

Language, Cognition, and Mind

James A. Hampton
Yoad Winter *Editors*

Compositionality and Concepts in Linguistics and Psychology

 Springer Open

Language, Cognition, and Mind

Volume 3

Series editor

Chungmin Lee, Seoul National University, Seoul, South Korea

Editorial board members

Tecumseh Fitch, University of Vienna, Vienna, Austria

Peter Gaerdenfors, Lund University, Lund, Sweden

Bart Geurts, Radboud University, Nijmegen, The Netherlands

Noah D. Goodman, Stanford University, Stanford, USA

Robert Ladd, University of Edinburgh, Edinburgh, UK

Dan Lassiter, Stanford University, Stanford, USA

Edouard Machery, Pittsburgh University, Pittsburgh, USA

This series takes the current thinking on topics in linguistics from the theoretical level to validation through empirical and experimental research. The volumes published offer insights on research that combines linguistic perspectives from recently emerging experimental semantics and pragmatics as well as experimental syntax, phonology, and cross-linguistic psycholinguistics with cognitive science perspectives on linguistics, psychology, philosophy, artificial intelligence and neuroscience, and research into the mind, using all the various technical and critical methods available. The series also publishes cross-linguistic, cross-cultural studies that focus on finding variations and universals with cognitive validity. The peer reviewed edited volumes and monographs in this series inform the reader of the advances made through empirical and experimental research in the language-related cognitive science disciplines.

More information about this series at <http://www.springer.com/series/13376>

James A. Hampton · Yoad Winter
Editors

Compositionality and Concepts in Linguistics and Psychology



Editors

James A. Hampton
Department of Psychology
City, University of London
London
UK

Yoad Winter
Utrecht Institute of Linguistics OTS
Utrecht University
Utrecht
The Netherlands



ISSN 2364-4109

Language, Cognition, and Mind

ISBN 978-3-319-45975-2

DOI 10.1007/978-3-319-45977-6

ISSN 2364-4117 (electronic)

ISBN 978-3-319-45977-6 (eBook)

Library of Congress Control Number: 2017941541

This book was advertised with the name of the publisher as copyright holder in error, whereas “The Editor(s) / The Author(s)” holds the copyright.

© The Editor(s) (if applicable) and The Author(s) 2017. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Acknowledgements

This volume grew out of a workshop on Concept Composition & Experimental Semantics/Pragmatics, which took place at Utrecht University in September 2013. That workshop and the work of the second editor were supported by a VICI grant 277-80-002 of the Netherlands Organisation for Scientific Research (NWO).

For their help with the reviewing process of the articles in this volume, the editors wish to thank Alan Bale, Peter Bosch, Oliver Bott, Heather Burnett, Zachary Estes, Nat Hansen, Lotte Hogeweg, Martin Jönsson, Dan Lassiter, Gregory Murphy, Marijn Struiksma and David Vinson, as well as the authors of the volume. We are also grateful to participants in the 2013 workshop whose works are not represented here but contributed a lot to the success of the workshop: Patricia Amaral, Marisa Casillas, Stephen Crain, Martin Everaert, Bart Geurts, Lotte Hogeweg, Hans Kamp, and, especially the late Remko Scha, whose wisdom and collegiality have inspired the work on this volume since its start.

Utrecht, The Netherlands
February 2017

Contents

Introduction	1
James A. Hampton and Yoad Winter	
Cognitively Plausible Theories of Concept Composition	9
Lawrence W. Barsalou	
Compositionality and Concepts—A Perspective from Formal Semantics and Philosophy of Language	31
Francis Jeffry Pelletier	
Compositionality and Concepts	95
James A. Hampton	
Typicality Knowledge and the Interpretation of Adjectives	123
Choonkyu Lee	
Concept Typicality and the Interpretation of Plural Predicate Conjunction	139
Eva B. Poortman	
Critical Typicality: Truth Judgements and Compositionality with Plurals and Other Gradable Concepts	163
Yoad Winter	
Complement Coercion as the Processing of Aspectual Verbs: Evidence from Self-paced Reading and fMRI	191
Yao-Ying Lai, Cheryl Lacadie, Todd Constable, Ashwini Deo and Maria Mercedes Piñango	
Conceptual Combination, Property Inclusion, and the Aristotelian-Thomistic View of Concepts	223
Christina L. Gagné, Thomas L. Spalding and Matthew Kostelecky	

**Conceptual Versus Referential Affordance
in Concept Composition** 245
Louise McNally and Gemma Boleda

**How Does the Left Anterior Temporal Lobe Contribute
to Conceptual Combination? Interdisciplinary Perspectives** 269
Masha Westerlund and Liina Pyykkänen

**Dimension Accessibility as a Predictor of Morphological
Gradability** 291
Galit W. Sassoon

Author Index. 327

Subject Index. 333

Contributors

Lawrence W. Barsalou Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, Scotland

Gemma Boleda Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain

Todd Constable Yale Magnetic Resonance Research Center, New Haven, USA

Ashwini Deo Department of Linguistics, The Ohio State, Columbus, USA

Christina L. Gagné Department of Psychology, University of Alberta, Edmonton, Canada

James A. Hampton Department of Psychology, City, University of London, London, UK

Matthew Kostelecky Department of Philosophy, University of Alberta, Edmonton, Canada

Cheryl Lacadie Yale Magnetic Resonance Research Center, New Haven, USA

Yao-Ying Lai Department of Linguistics, Yale University, New Haven, USA

Choonkyu Lee Department of Linguistics, University of Potsdam, Potsdam, Germany

Louise McNally Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain

Francis Jeffrey Pelletier Philosophy Department, University of Alberta, Edmonton, AB, Canada

Maria Mercedes Piñango Department of Linguistics, Yale University, New Haven, USA

Eva B. Poortman Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, The Netherlands

Liina Pykkänen Departments of Psychology and Linguistics, New York University, New York, USA; NYUAD Institute, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

Galit W. Sassoon Department of English Literature and Linguistics, Bar Ilan University, Ramat Gan, Israel

Thomas L. Spalding Department of Psychology, University of Alberta, Edmonton, Canada

Masha Westerlund Department of Psychology, New York University, New York, USA

Yoad Winter Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, The Netherlands

Introduction

James A. Hampton and Yoad Winter

Abstract By highlighting relations between experimental and theoretical work, this volume explores new ways of addressing the problem of concept composition, which is one of the central challenges in the study of language and cognition. An introductory chapter lays out the background to the problem. The subsequent chapters by leading scholars and younger researchers in psychology, linguistics and philosophy, aim to explain how meanings of different complex expressions are derived from simple lexical concepts, and to analyze how these meanings connect to concept representations. This work demonstrates an important advance in the interdisciplinary study of concept composition, where points of convergence between cognitive psychology, linguistics and philosophy emerge and lead to new findings and theoretical insights.

In every day of our conscious life, we are constantly busy manipulating abstract representations of the world around us. This human dexterity with mental representations, or *concepts*, has fascinated scholars since antiquity. Modern cognitive scientists formulate this fascination as two fundamental questions about the relations between language and thought. One question concerns the representation and organization in the mind and brain of concepts for words like *cat*, *big*, *sleep*, and *nearby*. How do these words refer to the external world? How do our minds represent their meaning? Another question is about how concepts interact and combine when humans use complex expressions like “the big cat that is sleeping nearby”. How do the meanings of individual words combine to create a sentence meaning, or a *proposition*? These two questions underlie studies of *simple concepts* and *concept composition*. This distinction between two problems is convenient, but it also reflects substantial gaps between current analyses of language understanding.

J.A. Hampton (✉)
Department of Psychology, City, University of London, London, UK
e-mail: j.a.hampton@city.ac.uk

Y. Winter
Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, The Netherlands
e-mail: y.winter@uu.nl

Since the 1970s, cognitive psychologists and theoretical linguists have taken rather different routes in their study of concepts and how they compose. In Cognitive Psychology, much research on concepts has been framed by the work of Eleanor Rosch (1975), which focused on simple words and the everyday concepts they denote. By contrast, in Theoretical Linguistics, and especially in formal and generative paradigms influenced by Montague (1973), the focus has been different. In this framework, research on natural language semantics has predominantly addressed the compositionality of meanings in relation to syntactic theory, with special emphasis on a few logical concepts like *every*, *some* or *most*.

This analytical gap between Psychology and Linguistics is undesirable, as was acknowledged early on in the history of natural language semantics (Katz and Fodor 1963): psychological theories cannot ignore concept composition, and natural language semantics cannot ignore the complex mental aspects of meaning. Later on, works in Cognitive Psychology following Osherson and Smith (1981), Cohen and Murphy (1984) and Hampton (1987) have started to explore hard problems of concept composition, especially in relation to short nominal expressions. The importance of this line of work for formal semantics has been stressed in Kamp and Partee (1995). More recently, and especially since the 2000s, more points of contact between formal semantics and various aspects of concepts have been developed in the study of lexical representations and vagueness in natural language (Pustejovsky 1995; McNally and Kennedy 2008; Asher 2011; Nouwen et al. 2011). Yet, despite similar motivations and related insights, much of the work on concepts and meaning composition has been carried out independently in Linguistics and Psychology.

This volume aims to help in bridging this gap. The articles here represent a substantial sample of the recent work done on concept composition in experimental semantics and psychology. The authors include leading authorities in their disciplines as well as younger researchers. Together, these works show the centrality and complexity of the concept composition problem, and the many ways in which it affects current research in Psychology, Linguistics and Philosophy.

The chapters by *Barsalou* and by *Pelletier* are broad overviews of concept composition from a psychological perspective and a linguistic-philosophical perspective, respectively. *Barsalou* starts from the observation that we use concepts in everyday life to understand our environment, solve problems and achieve goals. In this respect, it could be claimed that we have a common heritage with other intelligent forms of life, with a difference in degree of complexity but not in kind. From this perspective, language, a late-comer in the evolution of humans, is positioned on top of a highly developed and integrated system of intelligent conceptual thinking. All human languages have the same basic structure. There are atomic units (words or lexemes) which are combined according to syntactic rules (grammar) into phrases and sentences. Barsalou stresses that in addition to such “formal” considerations, situations, pragmatics, and imagery have all been found to play a role in how language utterances acquire meaning. Consequently, language meaning could not be effectively studied without also studying its users. Barsalou proposes that a concept can act as a tool for simulating experience. “John stroked

the cat” leads to an imaginative construction of a scene in which previous visual and tactile experiences are recruited to fill out the meaning of the sentence. This process explains how the verb can change depending on the context—stroking one’s beard, or stroking a putt in golf capture something of the same meaning (a gentle movement in contact with an object) but at the same time the situation simulated leads to many differences. Barsalou’s point is that to represent the meaning of “stroke” as a single lexical symbol STROKE (Agent, Patient) fails to account for the different inferences that are to be drawn from the simulation. Hence the meaning of the word “stroke”, for example its propensity to allow us to think true and false thoughts involving acts of stroking, can only be determined relative to its context of use, which triggers complex interactions with the simulated situations.

Pelletier argues the case for the more formal approach to semantics and meaning. He provides a “natural history” of different positions held by philosophers of language, linguists of different schools, and cognitive psychologists. He draws attention to a critical question that one should ask about any utterance, namely “what was really said?” The difficulty that Pelletier points to is one that has long been a contentious issue between philosophers and psychologists. Can the mental representation of a concept be treated as being the “real meaning” of a term? Aren’t there external aspects of meaning that are not inside the head?

If there are, as Pelletier argues (see also Rey 1983), how is this externalist view of semantics connected to the efforts of cognitive scientists to provide a “micro” account of how individual meanings arise with particular speakers and hearers and particular contextual situations? How is it possible to integrate the important and valuable insights into how different syntactic forms construct meanings for sentences with the empirical evidence of how speakers themselves construct meanings in context-dependent ways?

In the chapter by *Hampton*, these issues are given a further airing. Hampton attempts to answer the conundrum of how concepts can combine to meet the constraints of compositionality in the case where those concepts are in fact vague or fuzzy prototypes. He outlines the Composite Prototype Model for combining prototype concepts. Concepts are seen as being constituted by frame representations containing features expressed as dimensions and values (e.g. COLOUR = [RED, ORANGE]). When people have to form a conjunction or disjunction of prototype concepts, the model provides an account of the processes by which the two frames are amalgamated into a single composite frame. Along the way some properties of the concepts are lost, while new emergent properties may also be added to the composite, either from knowledge of the world or from the need to resolve incompatibilities between the two concepts. A key result supporting the model is the finding that, because of the merging of the two sets of prototypical features, noun phrases that appear to express a conjunction (e.g. “sports which are also games”) do not in fact receive a strict conjunctive interpretation when people categorize different exemplars. In a second major section, Hampton explains how this result relates to the distinction between two important theoretical notions linked to prototype theory, namely Typicality and Graded Truth. These two notions are then further differentiated from lack of knowledge or ignorance. Hampton

concludes his chapter by defending the Internalist/Descriptivist theory of conceptual contents. Many concepts, he argues, (and particularly those that are not “natural kinds” like snakes or lemons), are constituted by a structured set of intensional properties. Because we understand concepts this way we are free to generate fanciful and creative new ideas, unconstrained by their need to provide reference or determine truth in the real world.

The three chapters by **Lee**, **Poortman** and **Winter** all address questions about truth-value judgements. **Lee** reports results of two new experiments. One experiment shows a natural tendency to prefer certain colors and materials to others with different concepts. This preference is related to truth-value judgements with complex nominals. For instance, 71% of the participants in this experiment accepted the description *wooden bike* for a picture showing a bike that is only partially made of wood. By contrast, only 42% of the participants accepted the description *wooden frame* for a frame that contains the same amount of wood. This shows an important effect of typicality on truth-value judgements, in agreement with previous results about the correlations between typicality and membership, also described in Hampton’s chapter.

Poortman shows more new results about the relations between typicality and truth-value judgements. As Poortman’s experiments demonstrate, participants’ truth-value judgements on plural sentences like “the men are walking and writing” show sensitivity to the (attested) fact that walking and writing do not typically occur at the same time. Poortman shows that this atypicality increases tolerance to “split situations”, with two men walking and two men writing. Such split situations are less acceptable for sentences like “the men are walking and singing”, where there is no incompatibility between concepts.

Winter uses the results of Lee and Poortman’s experiments in a proposal for a general procedure of concept composition in formal semantics. As Winter stresses, truth-value judgements are not simply a matter of how odd or commonplace different situations are. Situations where two men are walking, and two other men are singing, are not odd at all, yet in Poortman’s experiment they are often rejected for sentences like “the men are walking and singing”. In Winter’s account, the tendency to accept sentences is boosted by *potential atypicality* of alternative situations. This has much in common with other new results about *reciprocal sentences* like “the girls are pinching each other”. Winter proposes an account that unifies the analysis of these results with the analysis of Lee’s and Poortman’s data.

The three chapters by **Lai et al.**, **Gagné et al.** and **McNally and Boleda** all address the processes that add to the understanding of complex concepts, beyond what is directly said. **Lai et al.** analyze verb phrase concepts in sentences like “*John began the book*”. In such sentences, concept composition may require understanding John’s relevant activity in relation to the book (reading, writing etc.). Lai et al. compare such sentences with sentences containing psychological verbs “*John enjoyed/disliked the book*” which also allow understanding of a relevant activity, but through presumably different linguistic mechanisms. Based on a self-paced reading and an fMRI study, Lai et al. argue against previous analyses, where relevant entities for events like reading and writing are added in the process

of semantic composition. Instead, they test the proposal that verbs like *begin* (but not *enjoy*) involve concepts that require *structured individuals*. When describing such individuals, Lai et al. rely on the understood internal structure of books and similar objects. As they suggest, such internal structures can be established along spatial (pages in a book), informational (e.g. the informational content of a book) or other dimensions. They conclude that concept composition involves construing this structure out of the complement triggered by the meaning requirements of the verb.

Gagné et al.'s chapter is in two parts. In the first they discuss their research into a recently discovered phenomenon known as the Modification (or Modifier) Effect and originally reported by Connolly et al. (2007). When a generic statement such as “candles are made of wax” is changed by adding an adjectival modifier to the subject noun (e.g. “purple candles are made of wax”) then people consider it less likely to be true, even though they have no reason to connect the modifier to the property. Spalding and Gagné (2015) have also shown an inverse effect—if people consider a generic statement to be false, then the modified version becomes more likely in people's estimation. Gagné et al. discuss how this phenomenon is a part of a theoretically influential and important research area concerning the combination of concepts to construct noun phrases, either with adjectival modification (*red shirt*) or through noun + noun compounding (e.g. *football shirt*). The topic has been extensively researched since it was first thrown out as a challenge to prototype theory in the seminal article by Osherson and Smith (1981). In the second part of their chapter, Gagné et al. outline a different conception of concepts based on Aristotelian-Thomistic (AT) philosophy. This framework for understanding concepts forms an intriguing bridge between the externalist semantics defended in large part by Pelletier and the embodied cognition approach propounded by Barsalou.

Like the chapters by Lai et al. and Gagné et al., the chapter by **McNally and Boleda** focuses on the processes that add to our understanding of complex concepts. From a theoretical and computational perspective, McNally and Boleda propose two processes that affect concept composition. One process, which they call *conceptual affordance*, allows humans to use information within the composed concepts themselves. For instance, one understanding of the complex expression *red box* uses the information contained within the simple concepts *red* and *box*. This process leads to the interpretation “box which is colored red”. By contrast, in another process, called *referential affordance*, information about the referent described by the phrase is used to guide the way in which the concepts in question are composed. This process also allows other interpretations, like “box which contains red objects” for the phrase *red box*. McNally and Boleda argue for a mixed semantic model for composition using the two kinds of affordance, and propose an integrated account that combines a formal semantic framework with distributional semantics, a computational approach that models conceptual aspects of word meaning.

Westerlund and Pykkänen present an account of a particular cortical region, the Left Anterior Temporal Lobe (LATL) which appears to have a direct role in many of the conceptual modification and combination processes already discussed in other chapters in the volume. Using Magnetoencephalography (MEG) they show

how this part of the cortex responds to conceptual combination. The greater the degree to which the modifier leads to specialization of the head noun, the more powerful is the response in the LATL. Westerlund and Pyllkkänen discuss different accounts of conceptual combination—in particular schema based models that lead to modification of the head noun schema (e.g. Smith et al. 1988), and relation based models that propose that meaning of a compound is generated by seeking an appropriate semantic relation to link the two nouns (e.g. a *dog house* is a house built for a dog). This work is connected with the chapters by Hampton, and by Gagné et al., and the neurological data adds valuable criteria for how different accounts may be evaluated.

Sassoon addresses a systematic difference between the way concepts are expressed by adjectives and nouns, which surfaces in the acceptability of their composition with *comparative expressions*. Many adjectives show their gradability with comparative forms like *richer*, *more expensive* or *more American*. By contrast, nouns do not easily allow constructions like *more a millionaire*, *more an American*, or *more a duck*. In agreement with the prototype theory of concepts, Sassoon proposes that both noun concepts and adjective concepts involve gradability. However, according to her proposal, only adjectives have their gradable aspects open for modification by comparative operations (e.g. *more* or *-er*). Sassoon describes some empirical and preliminary experimental results supporting her views on the question of gradability and concept composition.

References

- Asher, N. (2011). *Lexical meaning in context*. Cambridge University Press.
- Cohen, B., & Murphy, G. L. (1984). Models of concepts. *Cognitive Science*, 8(1), 27–58.
- Connolly, A. C., Fodor, J. A., Gleitman, L. R., & Gleitman, H. (2007). Why stereotypes don't even make good defaults. *Cognition*, 103(1), 1–22.
- Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, 15(1), 55–71.
- Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, 57, 129–191.
- Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, 39(2), 170–210.
- McNally, L., & Kennedy, C. (Eds.). (2008). *Adjectives and Adverbs: Syntax, semantics, and discourse*. Oxford: Oxford University Press.
- Montague, R. (1973). The proper treatment of quantification in ordinary English. In J. Hintikka, J. Moravcsik & P. Suppes (Eds.), *Approaches to Natural Languages: proceedings of the 1970 Stanford workshop on grammar and semantics*, D. Reidel, Dordrecht. Reprinted in R. Thomason, editor (1974), *Formal Philosophy: selected papers of Richard Montague*. New Haven: Yale.
- Nouwen, R., van Rooij, R., Sauerland, U., & Schmitz, H. C. (Eds.). (2011). *Vagueness in Communication*. Dordrecht: Springer.
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9, 35–58.
- Pustejovsky, J. (1995). *The generative lexicon*. Cambridge: MIT Press.
- Rey, G. (1983). Concepts and stereotypes. *Cognition*, 15(1), 237–262.

- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104(3), 192–233.
- Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). “Combining prototypes: A selective modification model”. *Cogn Sci*, 12(4), 485–527.
- Spalding, T. L., & Gagné, C. L. (2015). Property attribution in combined concepts. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 41(3), 693–707.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Cognitively Plausible Theories of Concept Composition

Lawrence W. Barsalou

Abstract If a theory of concept composition aspires to psychological plausibility, it may first need to address several preliminary issues associated with naturally occurring human concepts: content variability, multiple representational forms, and pragmatic constraints. Not only do these issues constitute a significant challenge for explaining individual concepts, they pose an even more formidable challenge for explaining concept compositions. How do concepts combine as their content changes, as different representational forms become active, and as pragmatic constraints shape processing? Arguably, concepts are most ubiquitous and important in compositions, relative to when they occur in isolation. Furthermore, entering into compositions may play central roles in producing the changes in content, form, and pragmatic relevance observed for individual concepts. Developing a theory of concept composition that embraces and illuminates these issues would not only constitute a significant contribution to the study of concepts, it would provide insight into the nature of human cognition.

The human ability to construct and combine concepts is prolific. On the one hand, people acquire tens of thousands of concepts for diverse categories of settings, agents, objects, actions, mental states, bodily states, properties, relations, and so forth. On the other, people combine these concepts to construct infinite numbers of more complex concepts, as the open-ended phrases, sentences, and texts that humans produce effortlessly and ubiquitously illustrate. Major changes in the brain, the emergence of language, and new capacities for social cognition all probably played central roles in the evolution of these impressive conceptual abilities (e.g., Deacon 1997; Donald 1993; Tomasello 2009).

In psychology alone, much research addresses human concepts (e.g., Barsalou 2012; Murphy 2002; Smith and Medin 1981) and concept composition (often referred to as *conceptual combination*; e.g., Costello and Keane 2000; Gagné and Spalding 2014; Hampton 1997; Hampton and Jönsson 2012; Medin and Shoben 1988; Murphy

L.W. Barsalou (✉)

Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, Scotland
e-mail: lawrence.barsalou@glasgow.ac.uk

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_2

1988; Wisniewski 1997; Wu and Barsalou 2009). More generally across the cognitive sciences, much additional research addresses concepts and the broader construct of compositionality (for a recent collection, see Werning et al. 2012).

1 Background Framework

A grounded approach to concepts. Here I assume that a concept is a dynamical distributed network in the brain coupled with a category in the environment or experience, with this network guiding situated interactions with the category's instances (for further detail, see Barsalou 2003b, 2009, 2012, 2016a, 2016b). The concept of *bicycle*, for example, represents and guides interactions with the category of bicycles in the world. Across interactions with a category's instances, a concept develops in memory by aggregating information from perception, action, and internal states. Thus, the concept of *bicycle* develops from aggregating multimodal information related to bicycles across the situations in which they are experienced. As a consequence of using selective attention to extract information relevant to the concept of *bicycle* from the current situation (e.g., a perceived bicycle), and then using integration mechanisms to integrate it with other bicycle information already in memory, aggregate information for the category develops continually (Barsalou 1999). As described later, however, background situational knowledge is also captured that plays important roles in conceptual processing (Barsalou 2016b, 2003b; Yeh and Barsalou 2006). Although learning plays central roles in establishing concepts, genetic and epigenetic processes constrain the features that can be represented for a concept, and also their integration in the brain's association areas (e.g., Simmons and Barsalou 2003). For example, biologically-based neural circuits may anticipate the conceptual structure of evolutionarily important concepts, such as agents, minds, animals, foods, and tools.

Once the conceptual system is in place, it supports virtually all other forms of cognitive activity, both online in the current situation and offline when representing the world in language, memory, and thought (e.g., Barsalou 2012, 2016a, 2016b). From the perspective developed here, when conceptual knowledge is needed for a task, concepts produce situation-specific simulations of the relevant category dynamically, where a simulation attempts to reenact the kind of neural and bodily states associated with processing the category. On needing conceptual knowledge about *bicycles*, for example, a small subset of the distributed bicycle network in the brain becomes active to simulate what it would be like to interact with an actual bicycle. This multimodal simulation provides anticipatory inferences about what is likely to be perceived further for the bicycle in the current situation, how to interact with it effectively, and what sorts of internal states might result (Barsalou 2009). The specific bicycle simulation that becomes active is one of infinitely many simulations that could be constructed dynamically from the bicycle network—the entire network never becomes fully active. Typically, simulations remain unconscious, at least to a large extent, while causally influencing cognition, affect, and

action. To the extent that part of a simulation becomes conscious, mental imagery is experienced. Such simulations need not provide complete or accurate representations, but are likely to be incomplete and distorted, representing abstractions, caricatures, and ideals, as well as specific learning episodes.

Situating concepts. As Pelletier (2017) notes, theories of concepts often focus either on concepts in the world from the formalist perspective (e.g., truth conditions) or on concepts in the brain from the cognitivist perspective (what Pelletier refers to as *objectivist* vs. *subjectivist* approaches). Although the grounded approach to concepts makes many assumptions about concepts from the cognitivist perspective, it additionally assumes that physical situations play central roles in establishing and using concepts (e.g., Barsalou 2003b, 2016b; Barsalou et al. 1993, 2007). From the grounded perspective, concepts are typically situated, that is, they become active to process some aspect of the current physical situation (or a physical situation in the past or future). Although concepts may sometimes become active independently of a situation, they typically become active to support effective action in a specific situation. As a result, concepts become coupled with their physical referents. Although grounded theories of concepts have a long way to go in developing satisfactory accounts of this coupling, they naturally anticipate it, and provide many mechanisms for understanding and explaining it. Thus, grounded theories offer an approach for unifying formalist and cognitivist accounts of concepts, and further assume that neither approach alone is sufficient.

Overview. The grounded account of concepts just described is not all that unusual given the diverse theories of conceptual processing that exist within the prototype, exemplar, connectionist, Bayesian, psycholinguistic, and grounded traditions of cognitive science (e.g., Asher 2011; Barsalou 1990; Bosch 1983; Casasanto and Lupyan 2015; Connell and Lynott 2014; Evans 2009; Jones and Love 2011; Lebois, Wilson-Mendenhall and Barsalou 2015; McClelland and Rumelhart 1985; Minda and Smith 2011; Pothos and Wills 2011). Many cognitively plausible theories assume increasingly that concepts are represented dynamically and are adapted to specific situations.

In contrast, theories of concept composition often assume relatively idealized accounts of concepts, where much of the dynamically varying content of concepts is omitted, given its complexity and messiness, focusing instead on a few key features and relations (e.g., some of the chapters in Werning et al. 2012). Even in the cognitivist accounts of concept composition mentioned earlier, concepts are often relatively idealized. Idealizing concepts has made establishing principles of concept composition tractable, and doing so has clearly been an effective strategy.

Nevertheless, avoiding the complexities of conceptual processing may obscure important principles of how concepts combine. As theories of concept composition mature increasingly, theorists and researchers appear willing to take on the challenge of dealing with these complexities (as illustrated by many of the chapters in this volume and also in Werning et al. 2012). In this spirit, I focus on three principles of human conceptual processing that strike me as central to its character: content variability, multiple representational forms, and pragmatic constraint. On

the one hand, these principles offer significant challenges to theories of concept composition. On the other, seriously addressing and incorporating these challenges may yield new theories that provide increasingly powerful and insightful accounts.

1.1 Content Variability

Lack of conceptual cores. Researchers often assume that a concept has a *core*, namely, important information about the respective category that is consistently activated rapidly and automatically, independently of context. Although cores are sometimes assumed to contain rules or definitions (e.g., “unmarried adult male human” for *bachelors*), they can also contain typical features that are neither necessary nor sufficient for category membership (e.g., “small flying feathered singing animal” for *birds*). Not only have traditional accounts of concepts proposed that concepts have cores (e.g., Barsalou 1982, 1989; Conrad 1978; Greenspan 1986; Whitney et al. 1985), so have current accounts (e.g., Dove 2009; Machery 2007; Mahon and Caramazza 2008).

An interesting wrinkle to core theories is whether cores become active early versus late (Lebois et al. 2015). According to most core theories, including those just cited, a conceptual core is the first information active for a concept, given its importance in representing the respective category (i.e., “core early” theories). Interestingly, however, major accounts of concepts over the years have assumed the opposite, namely, conceptual cores are only activated late, if and when needed (i.e., “core late” theories). Rather than being automatic and context-dependent, cores tend to be slow and optional.

In some core late theories, definitional information is considered relatively specialized and only necessary when careful judgments about a category must be made (e.g., Keil and Batterman 1984; Malt 1994; Smith et al. 1974). According to these theories, typical non-core information is sufficient most of the time for processing a category. When processing *water*, for example, it is usually sufficient to represent it as a colorless, clear, liquid used for daily needs, and widely available from faucets. Only on rare occasions does it become necessary to access and use core definitional information that water is H₂O (Malt 1994).

In other core late theories, core information tends to be active late because typical information is easier to activate, thus becoming active initially (e.g., Blanchette and Dunbar 2000; Forbus Gentner and Law 1995; Gentner et al. 1993; Ross 1984, 1987). From this perspective, cues first retrieve typical information from concepts related to superficial properties of the situation. Even though deeper knowledge might be available and relevant, it tends to be displaced by relatively superficial information that becomes active more easily. Only when strategic efforts to activate deeper, more important information are made does this latter information become available. Importantly for our purposes, however, core information in core late theories is not activated quickly and automatically across contexts.

Even more problematically for core early views, many theories argue that concepts don't have cores at all (e.g., Casasanto and Lupyan 2015; Connell and Lynott 2014; Evans 2009; McCloskey and Glucksberg 1979). According to these theories, all information in a concept is context-dependent, at least to some extent. Although some information for a concept may be highly entrenched (e.g., feathers for *birds*), it is still modulated by context. Indeed much recent research suggests that central information in concepts isn't activated automatically across contexts (for reviews, see Gawronski and Cesario 2013; Kiefer et al. 2012; Lebois et al. 2015; but see Augustinova and Ferrand 2014). For example, in the Stroop task that has long been believed to demonstrate the automatic activation of color features, it turns out that the accessibility of these features varies across contexts. In the Simon task that has long been believed to demonstrate the automatic activation of spatial features, it turns out that the availability of these features varies across contexts as well. In general, context modulates the activation of many kinds of central features across diverse forms of conceptual knowledge. Given the ubiquitous effects of contexts on the central features of concepts, it appears increasingly likely that the activation of all conceptual knowledge is at least somewhat context-dependent.

For our purposes here, such findings suggest that the concepts combined during concept composition typically vary widely in content (also see Asher 2011; Bosch 1983). When a given concept is combined with other concepts, its content is likely to vary considerably across concept compositions. Even when the same two concepts are combined on different occasions, their individual content is likely to vary, especially when combined in different background situations.

Situation-specific conceptual content. A related phenomenon is that the varying content of a concept often appears to contain information relevant in the current situation. In some of these cases, the information appears to be stored in the concept, with the current situation selecting it as relevant (e.g., Barsalou 1982; Conrad 1978; Greenspan 1986; Whitney et al. 1985). When processing the concept of *water* in the context of a Norwegian pond during winter, the property of *freezes* becomes active, although it would not become active normally (e.g., in the context of drinking *water* in a Norwegian restaurant). In other cases, varying content appears to originate in background situations (e.g., Barsalou and Wiemer-Hastings, 2005; Papies 2013; Wu and Barsalou 2009). When conceptualizing *chair*, for example, people might include chairs as occurring in offices. Similarly, when conceptualizing *truth*, people might represent truth as it occurs in courts of law. In general, considerable amounts of information in conceptual representations appear to originate in the background situation in which a concept is currently represented, including the setting, related objects, other agents, actions, and variety of internal states (e.g., goals, evaluations, affect, motivation, reward, mentalizing).

Furthermore, the background situations active for a given category are likely to vary widely, both between and within individuals. The concept of *chair*, for example, is likely to be associated in memory with background situations for kitchens, living rooms, classrooms, theaters, jets, and ski lifts. Not only do different kinds of chairs (and chair properties) become active when representing chairs in these different background situations, so does differing knowledge about the

external environment, action, and internal states. Sitting in a jet chair, for example, is associated with a different setting, actions, and internal states than sitting in a living room chair. Yeh and Barsalou (2006) review research in episodic and semantic memory showing that a given concept is often associated with a variety of background situations in memory. As a consequence, when a concept is processed in a particular situation, the situation activates related background information that enters into task performance.

Bayesian sampling. As we have just seen, the properties active for a concept vary with context. As we also saw earlier, concepts may not be associated with stable cores that are activated automatically. How, then, should we think about the varying content of a concept?

One possibility is that conceptual content is sampled in a Bayesian manner; in other words, sampling reflects both frequency of use and contextual relevance (Barsalou 2011). In a given situation, information that has been processed frequently across situations for the concept has a higher probability of becoming active than information processed infrequently (e.g., round vs. floats for *basketball*). Importantly, however, high probability information need not be core information, becoming active on all occasions when the concept is processed. Rather than being active obligatorily across situations, this information simply has a higher probability of being active than less frequent information. Depending on the context, high probability information may or may not be active (Lebois et al. 2015).

Additionally, information that is relevant for the concept in the current situation is more likely to become active than information less relevant in the situation. Thus, floats has a higher probability of becoming active for *basketball* in the context of needing a life preserver than in the context of playing pickup basketball. Putting these two factors together, the information that becomes sampled for a concept on a given occasion is a Bayesian function of its overall accessibility in memory and its situational relevance. Future research is needed to assess this conjecture, but given the general relevance of Bayesian processes in intelligent systems, Bayesian processes are likely to play central roles in conceptual processing as well.

General concepts. If the conceptual content represented for a concept varies dynamically across situations, how can a concept be represented in a general manner (cf. Pelletier, 2017)? The grounded approach to concepts offers several potential solutions to this issue.

First, the distributed network in the brain that aggregates multimodal information for a concept across its exemplars offers a general representation of the respective category. Because the network aggregates a tremendous amount of information across experiences with the category, it doesn't simply represent a single exemplar, but represents all exemplars together (see the distinction between simulations and simulators in Barsalou 1999, 2009, also 2016b). Although only one specific representation of the concept is constructed at a time dynamically from the network, the general concept is represented implicitly across the entire network. To the extent that people are aware of possessing much more knowledge about a concept than they can access in a given moment, they may believe that they possess a general representation at a meta-cognitive level. Furthermore, the entire network may enter

implicitly and unconsciously into cognitive functions, thereby producing general effects of the concept.

Second, the most accessible simulation that can be easily constructed from the network offers a default representation of the concept (cf. McNally and Boleda, 2017). Although the most accessible simulation is a specific simulation (or perhaps a family of closely-related simulations), it nevertheless offers a prototypical representation of all the simulations that the network is capable of producing. As a result, it potentially functions as a general representation of all possible simulations, in the spirit of global attractors (cf. McClelland and Rumelhart 1985) and prototype theory (cf. Hampton 2006).

Third, people may learn to explicitly construct specific simulations of a concept that offer a generic representation of it (e.g., one that omits situational detail). To the extent that people understand meta-cognitively that a concept can be general, they may be able to construct a simulation that captures this generic character. Notably, such simulations are simply some of the infinitely many simulations that the distributed multimodal network for a concept can construct. As a consequence, generic simulations reside together with situation-specific simulations that support situated action across diverse contexts. In a sense, generic simulations, too, are situation-specific simulations, with the relevant situations being those where generic representations are useful.

Sharing concepts. If the conceptual content represented for a concept varies dynamically, how can individuals ever come to effectively share a common conceptual representation in a specific situation (cf. Pelletier, 2017)? One likely solution to this problem follows from the facts that different individuals have similar bodies, brains, and cognitive systems; they live in similar physical environments; they operate in highly-coordinated social contexts. As a result, different individuals acquire similar distributed networks for a given concept over the course of development. Within a particular social group or culture, different individuals' networks are likely to be highly similar, given similar coordinated experiences with many shared exemplars. Even across different cultures, these networks are likely to be highly similar, given that all humans have similar bodies, brains, and cognitive systems, operating in similar physical and social environments.

As we have seen, however, the network for a concept is capable of producing an infinite number of dynamically-varying simulations across situations. How, then, do different individuals coordinate a shared simulation in a given situation, even when they share highly similar networks? One potential solution is that coordination results from establishing *common ground* (e.g., Clark 1996). Through the use of linguistic reference, shared experience, cultural convention, and pragmatic strategies, people coordinate their use of a concept in the current situation, attempting to construct similar simulations of it from their similar networks. As a result, different individuals may typically be able to establish a shared simulation relevant in the current situation, such that effective understanding and communication follows. Although the simulations that different conversants construct are never identical, they may typically be similar enough for communicative success (with ineffective communication occurring when the coordination of shared simulations fails).

Although much remains to be learned about how individuals coordinate dynamically constructed simulations, doing so seems feasible, given that establishing common ground and utilizing effective pragmatic strategies may effectively constrain the construction of simulations from shared conceptual networks.

Implications for concept composition. This section has reviewed three properties of concepts that present potential challenges for theories of concept composition. First, concepts typically do not appear to have stable cores. Second, the content active for a concept varies across situations. Third, the information active on a given situation for a concept may result from Bayesian sampling.

Together these three properties suggest that when a given concept is combined with others across situations, it is likely to exhibit no core stability and considerable situation-specific variability. Although representations of the concept across situations probably exhibit statistical regularities associated with Bayesian priors and situational relevance, they nevertheless take diverse forms, similar to how a given phoneme constantly adapts to its phonemic and articulatory contexts (e.g., Repp 1982). To the extent that these three properties are present in concepts, they are likely to play significant roles in concept composition, thereby placing challenging constraints on accounts of this process.

2 Multiple Representational Forms

Exemplars versus abstractions. Almost always, even psychological theories of concept composition assume that abstractions represent the concepts being combined, with these abstractions typically taking the form of rules, prototypes, and frames/schemata (e.g., Barsalou 1992, 2012; Barsalou and Hale 1993; Hampton 2006; Murphy 2002; Pothos and Wills 2011). Nevertheless, considerable evidence has accumulated that a given concept includes detailed information about specific category exemplars (e.g., Allen and Brooks 1991; Medin and Schaffer 1978; Nosofsky 2011; also see Minda and Smith 2011). Rather than abstracting over the idiosyncratic features of category exemplars, conceptual knowledge includes these features and correlations between them. Furthermore, when people later use this knowledge, they often appear to draw on exemplar-level information. Such findings have suggested to many researchers that memories of a category's exemplars represent a concept, not an abstraction.

It is worth noting, however, that abstractions (not just exemplars) are capable of representing detailed information about category exemplars (Barsalou 1990, 2003a; McClelland and Rumelhart 1985). Nevertheless, the important possibility remains that people often activate memories of specific exemplars when representing a concept. If so, then how does concept composition with the concept proceed? Even if an abstraction were to represent the concept, what are the implications for concept composition of exemplar-specific details that the concept probably contains?

Amodal symbols. Typically, accounts of concept composition assume that it operates on representations that are amodal, symbolic, and propositional (e.g.,

Barsalou 1999, 2008a, 2010). From this perspective, a symbolic representational structure, such as a relation, predicate, frame, or schema, has arguments, variables, or slots, whose values vary across instances and situations (e.g., Barsalou 1992). Given these kinds of representational structures, concept composition becomes readily feasible (e.g., Gagné and Spalding 2014; Wisniewski 1997). To implement the combination of an object with a property, for example, a variable in the frame for the object is set to a specific value. When representing the concept composition of *green flamingo*, for example, the *color* variable in the *flamingo* frame is set to the value *green*. Similarly, to implement the combination of two nouns, a frame assigns the two nouns as arguments of different variables within a relation, predicate, frame, or schema. When representing the concept composition of *flamingo fare*, a frame for *eat* could take *flamingo* and *fare* as values of the variables for *agent* and *theme*, respectively.

Other accounts of concept composition assume that feature sets for individual concepts are combined in varying ways to produce the meanings of complex phrases (e.g., Hampton 1997; Hampton and Jönsson 2012). Again, the respective features are typically assumed to be amodal symbolic representations. When representing *sports that are games*, for example, people combine the feature sets for *sports* and *games* to produce the combined concept.

Amodal representations such as frames and features may play central roles in the representation of concepts, and in turn, during the process of combining them. Nevertheless, other important representational processes appear to play important roles in concept composition as well, in particular, multimodal simulations and linguistic forms. If so, then theories of concept composition may benefit from taking them into account.

Multimodal simulations. Earlier, in the introduction, the construct of multimodal simulation was introduced when defining *concept*. To reiterate briefly, a distributed multimodal network for a category in the brain attempts to simulate the kind of neural and bodily states that occur while interacting with a category's members. When simulating a bicycle, for example, the brain and body reenact the kinds of states that occur while experiencing bicycles, including how they look, where they are used, how to ride one, and associated internal states. Considerable evidence has accumulated supporting the proposal that simulation is a basic computational process in the brain, not only in conceptual processing, but in all cognitive processes, ranging from perception to social cognition (e.g., Barsalou 2008a, 2016b). Certainly, multimodal simulation is not the only representational process in the brain, but it appears to be a central one used widely and frequently. If so, then it may contribute significantly to the process of combining concepts.

To date, little research has addressed the role of simulation in concept composition directly. One set of experiments, however, provides some initial evidence (Wu and Barsalou 2009). Across three experiments, a preliminary prediction was that when participants produced the features of an object, they would simulate a typical experience of the object, reporting features salient in the simulation. As a consequence, features occluded in the simulation, and therefore not salient, wouldn't be produced frequently. When participants produce features of a typical *lawn*, for example, they should simulate a green lawn. Because *green* and *grass*

blades are salient in the simulation, they should be produced frequently. Because *dirt* and *roots* are occluded, they should be produced rarely. Across experiments, salient features were produced much more often than occluded features, as the simulation account predicts.

The primary prediction across experiments was that when the same objects were combined with modifiers that revealed normally occluded features, these occluded features would become salient and therefore be reported more frequently. When participants received the concept composition of *rolled-up lawn*, for example, they should simulate a rolled-up lawn. Because *dirt* and *roots* become unoccluded in the simulation and therefore salient, they should be reported more often than for *lawn* alone. Across experiments, this finding occurred robustly. Not only did it occur for familiar phrases such as *half-watermelon*, it occurred for unfamiliar phrases such as *glass car*.

Most importantly, however, unoccluded features did not become more salient for noun phrases such as *rolled-up snake*. This finding suggests two conclusions. First, modifiers such as *rolled-up* do not have symbolic rules associated with them specifying that the occluded features of any noun they modify become salient. If such rules existed, then normally occluded features should always become more salient whenever the modifier combines with any noun. Again, however, occluded features did not become more salient for phrases such as *rolled-up snake*. Second, modifiers such as *rolled-up* only change the salience of normally occluded features when the process of rolling up an object unoccludes these features, as when someone rolls up a lawn, but not when a snake rolls up itself. Because it's unlikely that representations of *rolled-up* with every possible noun are already stored in memory, it's likely that most interpretations of these noun phrases are computed online via concept composition, as the phrase is heard. Although it is in principle possible to develop an amodal procedure that correctly increases (or doesn't increase) the salience of occluded features across all *rolled up* noun phrases, a simpler approach exists: Simply simulate rolling up the noun object, and report the features salient in the process, especially in the end state. Rather than relying on a complex amodal rule to determine the concept composition, a simulation of the noun phrase naturally specifies whether occluded features become salient or not.

More generally, simulation may play central roles across many kinds of concept composition. Consider Medin and Shoben's (1988) example of *gray hair* versus *gray cloud*, where *gray* takes on different values in the two noun phrases. Although amodal symbols could capture such values, a simpler and more intuitive account is that simulating *gray hair* and *gray cloud* lead to different shades of gray being simulated. Because people store memories of *gray hair* and *gray cloud*, they can easily retrieve and simulate these memories to produce meanings of the respective phrases. Similarly, consider the noun-noun combinations of *mountain trail* and *desert trail*. By retrieving and simulating memories of each, the resulting interpretations implicitly represent the features that *mountain trails* tend to change elevation and direction more often than do *desert trails*.

Finally, multimodal simulations capture affordances that may often occur during concept composition. The phrase *handleless cup*, for example, has the affordance of being grasped with a power grip rather than with a precision grip. On hearing this phrase in a sentence like, “Would you enjoy drinking from a handleless cup,” much work suggests that people would immediately produce a motor simulation of picking the cup up with a power grip to answer the question (e.g., Chaigneau et al. 2004; Heard et al. 2015). Because multimodal simulations carry affordances of objects and situations, activating these simulations during the concept composition that underlies language use prepares comprehenders for effective situated action in the world.

As these examples illustrate, multimodal simulation offers a natural means of establishing the integration of conceptual information that occurs in concept composition. Not only does multimodal simulation offer a means of combining concepts never combined before as just described (e.g., *handleless cup*), it also offers a means of explaining the extensional feedback that underlies the learned representations of many familiar phrases (e.g., *tea cup*), as described later. Multimodal simulation is also compatible with the graphical diagrams used widely across cognitive linguistics to represent the concept composition that underlies spatial relations, part-whole relations, and event structure (e.g., Langacker 1986, 1987, 2008; Talmy 1983, 1988, 2000).

Distributed linguistic representations. Still another potentially important representational mechanism for implementing concept composition is the distribution of linguistic forms associated with a concept (with linguistic forms including words, phrases, and sentences). Depending on the approach, a distributed linguistic representation can simply include the word associates of a concept, or it can be the surrounding words and syntax that typically occur in sentences that include the concept. Following Paivio’s (1971, 1986) Dual Coding Theory, many cognitive psychologists have proposed that linguistic forms work together with an image-based conceptual system to produce intelligent computation across a wide variety of cognitive tasks (for several recent perspectives, see Barsalou, Santos et al. 2008; Glaser 1992; Louwse 2008, 2011; Louwse and Connell 2011).

Other researchers from the tradition of distributional semantics have similarly suggested that distributed linguistic representations play a wide variety of important roles in conceptual processing (Baroni and Lenci 2010; Erk 2012; Erk and Padó 2008; Landauer and Dumais 1997; Landauer et al. 2013; Padó and Lapata 2007). Characterizing the word for a concept in terms of its distributed linguistic representation reflects the word’s semantics, such that its similarity to the semantics of other words can be established. To the extent that two concepts tend to cooccur with the same linguistic forms, the more similar they are semantically. As much research shows, the distributed linguistic representation associated with a concept affects its processing across a wide variety of tasks and phenomena (e.g., Barsalou et al. 2008; Erk 2012; Landauer et al. 2013; Louwse 2011). Thus, it wouldn’t be surprising if distributed linguistic representations enter into concept composition as well. As McNally and Boleda (2017) suggest, distributed linguistic representations could play central roles both in establishing both default interpretations of combined

concepts and situationally-relevant interpretations (consistent with the Bayesian account of conceptual variability presented earlier).

Consider two further possibilities. First, the distributed linguistic representations of two words being combined may suggest relations for combining the underlying concepts. When establishing the meaning of *bicycle corridor*, for example, the distributed linguistic representations for *bicycle* and *corridor* may be established initially. Because *walk* is a salient word associate of *corridor*, it becomes active quickly. Once *walk* becomes active for *corridor*, it then matches the word for *walk* that is stored with *bicycle* but is currently inactive (because *ride* is more highly associated with *bicycle* than is *walk*). As a result of this match, *walk* becomes active for *bicycle*, leading to the interpretation of *bicycle corridor* as a place where one walks a bicycle. Clearly, other relational interpretations are possible, but this interpretation illustrates how combining distributed linguistic representations could help converge one of the many possible concept compositions available.

A second possibility is that the distributed linguistic representations of two combined words project into long-term memory together, activating memories that contain words from both. Imagine reading a newspaper article in the past about a bicycle race around the underground snow corridors of a Canadian university, with the phrase *bicycle corridor* not being mentioned. When hearing the phrase *bicycle corridor* on a later occasion, the distributed linguistic representations generated from the two words could map onto words from the earlier text, activate the associated memory, and influence how *bicycle corridor* is currently understood. Rather than being interpreted as a corridor where bicycles are walked, *bicycle corridor* is now interpreted instead as a corridor where bicycles are raced.

Finally, researchers working from the distributed semantics perspective increasingly explore distributed representations of non-linguistic information, such as the images associated with a concept (e.g., Anderson et al. [in press](#); Thill et al. 2014). To the extent that the distributions of images associated with two concepts are similar, their conceptual content is likely to be similar as well. Interestingly, the distribution of images associated with a concept may belong to the distributed multimodal networks that represent concepts in grounded approaches (cf. Barsalou 1999, 2009, 2016a, 2016b).

Implications for concept composition. As this section has described, a concept can be associated with a wide variety of representational forms. A concept can be represented with abstractions, exemplars, or both. Orthogonally, these abstractions and exemplars could be represented with amodal symbols, multimodal simulations, distributed linguistic representations, or some combination. At this point, we don't really know which of these forms enter into concept composition. It wouldn't be surprising, though, if they all did, depending on the kinds of concepts being combined, together with the goals and constraints associated with a given processing context.

3 Pragmatic Constraint

Extensional feedback. Researchers argue increasingly that cognition cannot be studied independently of sensory-motor systems, the body, and the environment (e.g., Aydede and Robbins 2009; Barsalou 2016a, 2016b, 2008a, 2010; Barsalou et al. 2007; Clark 1998, 2008). From this perspective, it may be impossible to understand concept composition adequately without taking into account how it is grounded in these important ways.

Hampton's (1988) construct of *extensional feedback* illustrates how the environment can contribute to concept composition (see McNally and Boleda, 2017, for the related construct of *referentially-afforded modification*). Knowing that a *pet bird* is usually caged probably doesn't result from conceptual knowledge stored with *pet* or with *bird*, given that neither is usually caged. Instead, this knowledge originates in seeing that the pet birds in one's experience typically live in cages. Thus, the meaning of *pet bird* is established by activating memories of caged birds in the world—its meaning doesn't solely result from conceptually combining *pet* and *bird*.

The visual world paradigm developed by Tanenhaus et al. (1995) offers another example of extensional feedback (for a recent review, see Huettig et al. 2011). In this paradigm, participants perceive a visual world and are requested to act on it. For example, a participant might see objects laid out on a table and be asked to move one object on top of another. As participants hear various phrases, such as "apple on a towel," they have to figure out what these phrases mean. Much work demonstrates that the meanings of these phrases don't only result from concept composition; they also result from extensional feedback via the visual world. The visual world paradigm offers much potential for studying both concept composition and the role of extensional feedback.

Effects of extensional feedback are likely to vary across cultures, further demonstrating situational effects on concept composition. Imagine the classic example of a *pet fish*. For many westerners who have had goldfish and guppies as pet fish, the extensional feedback from these experiences probably contributes to how *pet fish* is interpreted. Imagine, though, a culture that doesn't have pet fish. What interpretations of *pet fish* would they produce? Without extensional feedback, they would have to rely on concept composition to construct interpretations, such that *pet fish* might well reflect a combination of what's typical for the two constituent concepts, *pet* and *fish* (e.g., Hampton 1988; Hampton and Jönsson 2012). Alternatively, imagine cultures that keep different kinds of pet fish than do westerners, for example, relatively wild fish kept in backyard ponds, or in walled-off areas of streams and rivers. For these cultures, extensional feedback from the respective experiences would produce still different interpretations of *pet fish*, again not based on concept composition but understood as a conventional phrase.

Multimodal simulation, discussed earlier, offers a natural mechanism for explaining how extensional feedback becomes established in memory. When experiencing a pet fish, for example, a multimodal sensory-motor representation of

the fish is established, including its background situation. If someone uses the phrase “pet fish” to describe the fish in this situation, the phrase becomes stored with the multimodal memory. On hearing the phrase, “pet fish,” on a later occasion, the associated memory becomes active as a simulation, which then represents the meaning of the phrase. In this way, the simple process of capturing and running multimodal simulations may contribute to the comprehension of concept compositions.

Finally, concept composition interacts with extensional feedback in perhaps unanticipated ways. Imagine never having heard the phrase, “pet fish” before and never having encountered a pet fish. On hearing someone say, “Meet my pet fish,” as they point to a goldfish in a glass tank, it is necessary to make sense of *pet fish* in this context. Even though an excellent example of a pet fish is readily perceptible in the situation, understanding the uttered phrase, *pet fish*, and mapping it onto its referent requires concept composition. Clearly, a variety of situational cues make this possible, but it is still probably necessary to combine *pet* and *fish* conceptually so that understanding what *pet fish* means can occur. This suggests that concept composition plays important roles in understanding immediate situations, and doesn’t only occur during language comprehension about non-present situations. Counter-intuitively, perhaps, understanding a phrase or sentence about a non-present situation may often draw heavily on extensional feedback, whereas understanding what a phrase or sentence means in the current situation may sometimes rely more on concept composition.

Situational constraint. As described earlier, people’s knowledge of a category includes knowledge of background situations in which its instances occur. Imagine hearing the sentence, “In my kitchen, I sat on an uncomfortable chair for three hours straight.” To comprehend the concept composition, *uncomfortable chair*, someone might represent the *chair* as a simple solid wooden chair of the type typically found in kitchens, and might represent *uncomfortable* as a sore bottom from sitting too long on hard wood. In contrast, imagine hearing, “On the jet, I sat in an uncomfortable chair for three hours straight.” Here someone might represent *chair* as a cramped jet chair, and *uncomfortable* as stiffness from not being able to move.

As these examples illustrate, the background situation constrains the meaning of a concept composition, such as *uncomfortable chair*, with it taking on different meanings in different situations (for a review, see Yeh and Barsalou 2006). With respect to the principle of variable content described earlier, situational constraint offers one central mechanism that causes conceptual content to change across situations. As people comprehend words, phrases, and sentences, they heavily utilize information in the background situation to construct meaning, a point well-established in the comprehension literature (e.g., Bransford and Johnson 1972, 1973). Although the principle of situational constraint is related to the earlier principle of situation-specific conceptual content, the two principles differ in emphasis. Whereas situation-specific conceptual content focuses on situational information as a source of varying conceptual content, situational constraint focuses on the pragmatic influence that the current situation has on establishing varying content.

Implications for concept composition. As this section has described, interactions with past and present situations appear to play central roles in how people combine concepts. Often, memories of experienced entities and events from past situations provide the meanings of current phrases and sentences via extensional feedback. Similarly, the relevant background situation constrains how concept composition proceeds, biasing the conceptual content activated and combined toward situationally relevant information. It is likely that many other forms of pragmatic constraint not anticipated here contribute to concept composition as well.

4 Conclusion

From the perspective of human cognition, principles of conceptual processing have potential implications for theories of concept composition. First, the content active for a concept on a given occasion appears to vary considerably. Rather than having a stable core, a concept exhibits dynamically varying content across situations, perhaps resulting from Bayesian sampling. Second, multiple representational forms may constitute the dynamically constructed representation of a concept on a given occasion. Although these representations could include abstractions, they may also include exemplars, and are at least likely to include exemplar-level information. Additionally, multimodal simulations and distributed linguistic representations may become active and play central roles in concept composition, along with amodal symbols. On a given occasion, the specific mix of representational forms may vary widely. Third, the dynamic representation of a concept on a given occasion is likely to reflect a variety of pragmatic constraints. On the one hand, these representations could contain conceptual content established through extensional feedback on earlier occasions. On the other, they could contain information associated with the current background situation.

The principles of conceptual processing just described may be incorrect. Much controversy remains concerning the nature of the human conceptual system, such that some researchers, at least, would take issue with the principles of conceptual processing presented here. Nevertheless, based on the current empirical literatures in cognitive science, it is likely that at least some of these principles are approximately correct, and that the human conceptual system does not take the idealized forms assumed in some accounts of concept composition.

A possible framework for understanding concept composition. Let's assume for the moment that the account of conceptual processing presented here is roughly correct. For the sake of argument, how might we explain concept composition if these principles are true? What kinds of theories might be successful under conditions of varying conceptual content, multiple representational forms, and pragmatic constraint?

Under these conditions, it seems difficult to imagine how a formal symbolic theory could be successful. Perhaps I'm being uncharitable, but a successful theory of this type would require that considerable information about exemplars,

situations, and multimodal content be formalized, as would be procedures for combining it (consider what would be required to handle the representations of *rolled-up lawn* vs. *rolled-up snake* described earlier).

Perhaps a more tractable approach is to assume that when people perform concept composition, they typically attempt to construct a multimodal simulation of a relevant situation (Barsalou 1999, 2008b; Wu and Barsalou 2009). In the process, people draw heavily on previous situational experience, as they attempt to construct a simulation that includes all the concepts being combined in a sensible manner. Certainly, various types of grammatical and linguistic patterns may contribute to this process. Nevertheless, attempting to construct a coherent simulation may be central.

Notably, simulation accounts of concept composition differ considerably from symbolic accounts. Whereas simulation accounts represent conceptual content implicitly in multimodal representations captured from situational experience, symbolic approaches abstract this information from situational experience and encoded it explicitly into amodal symbolic structures. As a consequence, the processes that underlie concept composition differ considerably at the level of cognitive and neural mechanisms (cf. Marr 1982). In simulation accounts, as we saw earlier for Wu and Barsalou (2009), complex simulations of experience are constructed and transformed, whereas in symbolic accounts, explicit symbolic structures are combined according to the rules of logic and language. Although, at an abstract level, the information being represented in the two approaches may be comparable, the representation and processing of this information differs considerably, leading to different explanations and predictions at the cognitive and neural levels.

Although it remains to be seen whether a simulation approach to concept composition will work, this approach has several qualities that might yield useful outcomes. First, the simulation approach relies heavily on situational experience, both past and current, consistent with the work reviewed earlier on the importance of situations for conceptual processing. Not only does the multimodal simulation constructed to represent a concept composition draw heavily on situational memories, it also draws pragmatically on the current physical situation.

Second, this approach seems well suited for handling variable conceptual content and multiple representational forms. From this perspective, it is not a problem that a concept's content varies constantly—it only matters whether its current content can be integrated with the evolving multimodal simulation of the relevant situation. Because the process of integrating simulations seems potentially “sloppy,” varying content might not constitute a severe problem. Similarly, it doesn't really matter whether an exemplar or an abstraction becomes active, because either could be integrated into an evolving simulation. Additionally, distributed linguistic representations could become active throughout the simulation process, activating and structuring relevant simulation content.

Third, this approach doesn't require a tight formalism to be effective. Rather than linguistic forms being associated with semantic formulae that combine in a compositional manner, linguistic forms and grammatical structures constitute instructions for constructing, transforming, and evaluating multimodal simulations (Barsalou 1999; cf. Langacker 1986, 1987, 2008). Because linguistic forms and

grammar become cues for building simulations, their formal properties are less important than the effects that they have on the simulation process.

Clearly, however, many approaches for explaining concept composition exist, and many new ones are likely to develop. Perhaps what is most significant and encouraging currently is that diverse groups of researchers increasingly attempt to understand concept composition from different perspectives (e.g., the contributors to the current volume and to Werning et al. 2012). Because concept composition may constitute an unusually challenging aspect of human cognition to explain, it seems likely that adopting different theoretical perspectives will be essential, as will be incorporating many kinds of empirical evidence.

Pelletier (2017) offers a long list of challenges for cognitivist approaches to concepts, noting that cognitivist approaches have yet to provide accounts of singular terms, quantification, grammatical categories, syntax, integrated thoughts, negation, and so forth. Clearly, these are important issues, and cognitive approaches have done relatively little to address them. To a large extent this is because cognitive approaches have focused on other issues, often attempting to establish the most basic mechanisms in the brain that underlie conceptual processing, holding off on more challenging issues until basic issues are better understood. Additionally, cognitive researchers often lack the technical expertise required to perform informed empirical study of more complex issues.

Rather than challenging the cognitivists to explain these issues on their own, perhaps formalists could work together with cognitivists to develop accounts that integrate both perspectives. Ultimately, optimal accounts of conceptual processing are likely to reflect sophistication about both its formal and cognitive aspects.

Furthermore, to the extent that the formal aspects of conceptual processing reflect structure in the world (e.g., truth conditions), grounded cognition offers a natural theoretical framework for integrating the formal and cognitive perspectives. Because grounded cognition focuses on coupling cognition with the environment, it sees the structure of concepts in the world as indispensable to conceptual processing (although it might take a more statistical view of the physical world as opposed to a logical one). From the grounded perspective, a successful account of conceptual processing is only possible once both perspectives are included and unified. A pure cognitive account on its own is far from sufficient, as is a pure formal account of structure in the world. Because cognition is coupled with the world, taking each seriously in explaining the other is essential.

References

- Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, 120, 3–19.
- Anderson, A., Bruni, E., Lopotolo, M., Poesio, M., & Baroni, M. (2015). Reading visually embodied meaning from the brain: Visually grounded computational models decode visual-object mental imagery induced by written text. *NeuroImage*, 120, 309–322.

- Asher, N. (2011). *Lexical meaning in context: A web of words*. Cambridge: Cambridge University Press.
- Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science*, 23, 343–348.
- Aydede, M., & Robbins, P. (2009). *The Cambridge handbook of situated cognition*. Cambridge: Cambridge University Press.
- Baroni, M., & Lenci, A. (2010). Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, 36, 673–721.
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, 10, 82–93.
- Barsalou, L. W. (1989). Intraconcept similarity and its implications for interconcept similarity. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 76–121). Cambridge: Cambridge University Press.
- Barsalou, L. W. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In T. K. Srull & R. S. W. Jr. (Eds.), *Content and process specificity in the effects of prior experiences: Advances in social cognition* (Vol. 3, pp. 61–88). Hillsdale, NJ: Erlbaum.
- Barsalou, L. W. (1992). Frames, concepts, and conceptual fields. In A. Lehrer & E. F. Kittay (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 21–74). Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou, L. W. (2003a). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 358, 1177–1187.
- Barsalou, L. W. (2003b). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, 18, 513–562.
- Barsalou, L. W. (2008a). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Barsalou, L. W. (2008b). Grounding symbolic operations in the brain's modal systems. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 9–42). New York: Cambridge University Press.
- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 1281–1289.
- Barsalou, L. W. (2010). Grounded cognition: past, present, and future. *Topics in Cognitive Science*, 2, 716–724.
- Barsalou, L. W. (2011). Integrating Bayesian analysis and mechanistic theories in grounded cognition. *Behavioral and Brain Sciences*, 34, 191–192.
- Barsalou, L. W. (2012). The human conceptual system. In M. Spivey, K. McRae, & M. F. Joannisse (Eds.), *The Cambridge handbook of psycholinguistics* (pp. 239–258). New York: Cambridge University Press.
- Barsalou, L. W. (2016a). Can cognition be reduced to action? Processes that mediate stimuli and responses make human action possible (pp. 81–96). In A. K. Engel, K. J. Friston, & D. Kragic, *Where's the action? The pragmatic turn in cognitive science (Strüngmann Forum Reports, Vol. 18. J. Lupp, Series Ed.)*. Cambridge, MA: MIT Press.
- Barsalou, L. W. (2016b). Situated conceptualization: Theory and application. In Y. Coello & M. H. Fischer, *Foundations of embodied cognition* (pp. 11–37). East Sussex: Psychology Press.
- Barsalou, L. W., Breazeal, C., & Smith, L. B. (2007). Cognition as coordinated non-cognition. *Cognitive Processing*, 8, 79–91.
- Barsalou, L. W., & Hale, C. (1993). Components of conceptual representation. From feature lists to recursive frames. In I. Van Mechelen, J. A. Hampton, R. Michalski, & P. Theuns, *Categories and concepts: Theoretical views and inductive data analysis* (pp. 97–144). San Diego: Academic Press.
- Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245–283). Oxford: Oxford University Press.

- Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. A. Zwaan, (pp. 129–163). New York: Cambridge University Press.
- Barsalou, L. W., Yeh, W., Luka, B. J., Olseth, K. L., Mix, K. S., & Wu, L.L. (1993). Concepts and meaning. In K. Beals, G. Cooke, D. Kathman, K. E. McCulloch, S. Kita, & D. Teste, *Chicago Linguistics Society 29: Papers from the parasession on conceptual representations* (pp. 23–61). University of Chicago: Chicago Linguistics Society.
- Blanchette, I., & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory & Cognition*, 28, 108–124.
- Bosch, P. (1983). “Vagueness” is context-dependence. A solution to the Sorites paradox. In T. T. Ballmer & M. Pinkal, *Approaching vagueness* (pp. 189–210). Amsterdam: North Holland.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717–726.
- Bransford, J. D., & Johnson, M. K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing* (pp. 383–438). Oxford: Academic Press.
- Casasanto, D., & Lupyan, G. (2015). All concepts are ad hoc concepts. In E. Margolis & S. Laurence, *The conceptual mind: New directions in the study of concepts* (pp. 543–566). Cambridge, MA: MIT Press.
- Chaigneau, S. E., Barsalou, L. W., & Sloman, S. A. (2004). Assessing the causal structure of function. *Journal of Experimental Psychology: General*, 133, 601–625.
- Clark, H.H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, A. (1998). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford: Oxford University Press.
- Connell, L., & Lynott, D. (2014). Principles of representation: Why you can’t represent the same concept twice. *Topics in Cognitive Science*, 6, 390–406.
- Conrad, C. (1978). Some factors involved in the recognition of words. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic factors in cognition* (pp. 103–120). Hillsdale, NJ: Erlbaum.
- Costello, F. J., & Keane, M. T. (2000). Efficient creativity: constraint-guided conceptual combination. *Cognitive Science*, 24, 299–349.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: WW Norton & Company.
- Donald, M. (1993). *Precis of Origins of the modern mind: Three stages in the evolution of culture and cognition*. *Behavioral and Brain Sciences*, 16, 737–748.
- Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, 110, 412–431.
- Erk, K. (2012). Vector space models of word meaning and phrase meaning: A survey. *Language and Linguistics Compass*, 6, 635–653.
- Erk, K., & Padó, S. (2008). A structured vector space model for word meaning in context. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp. 897–906). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from.
- Evans, V. (2009). *How words mean: Lexical concepts, cognitive models, and meaning construction*. Oxford: Oxford University Press.
- Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19, 141–205.
- Gagné, C. L., & Spalding, T. L. (2014). Conceptual composition: The role of relational competition in the comprehension of modifier-noun phrases and noun-noun compounds. *The Psychology of Learning and Motivation*, 59, 97–130.
- Gawronski, B., & Cesario, J. (2013). Of mice and men: What animal research can tell us about context effects on automatic responses in humans. *Personality and Social Psychology Review*, 17, 187–215.

- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, *25*, 524–575.
- Glaser, W. R. (1992). Picture naming. *Cognition*, *42*, 61–105.
- Greenspan, S. L. (1986). Semantic flexibility and referential specificity of concrete nouns. *Journal of Memory and Language*, *25*, 539–557.
- Hampton, J. A. (1988). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 12.
- Hampton, J. A. (1997). Conceptual combination. In K. Lamberts & D. R. Shanks (Eds.), *Knowledge, concepts, and categories* (pp. 133–159). East Sussex: Psychology Press.
- Hampton, J. A. (2006). Concepts as prototypes. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory (Vol 46)* (pp. 79–113). San Diego, CA, US: Elsevier Academic Press.
- Hampton, J. A., & Jönsson, M. (2012). Typicality and compositionality: The logic of combining vague concepts. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 385–482). Oxford: Oxford University Press.
- Heard, A. W., Masson, M. E., & Bub, D. N. (2015). Time course of action representations evoked during sentence comprehension. *Acta Psychologica*, *156*, 98–105.
- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*, 151–171.
- Jones, M., & Love, B. C. (2011). Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences*, *34*, 169–188.
- Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal Learning and Verbal Behavior*, *23*, 221–236.
- Kiefer, M., Adams, S. C., & Zovko, M. (2012). Attentional sensitization of unconscious visual processing: Top-down influences on masked priming. *Advances in Cognitive Psychology*, *8*, 50–61.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.
- Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2013). *Handbook of latent semantic analysis*. East Sussex: Psychology Press.
- Langacker, R. W. (1986). An introduction to cognitive grammar. *Cognitive Science*, *10*, 1–40.
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Lebois, L. A. M., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are automatic conceptual cores the Gold Standard of semantic processing? The context-dependence of spatial meaning in grounded congruency effects. *Cognitive Science*, *39*, 1764–1801.
- Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review*, *15*, 838–844.
- Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, *3*, 273–302.
- Louwerse, M. M., & Connell, L. (2011). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science*, *35*, 381–398.
- Machery, E. (2007). Concept empiricism: A methodological critique. *Cognition*, *104*, 19–46.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*, 59–70.
- Malt, B. C. (1994). Water is not H₂O. *Cognitive Psychology*, *27*, 41–70.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY: Henry Holt.

- McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159–188.
- McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, *11*, 1–37.
- McNally, L., & Boleda, G. (2017). Conceptual vs. referential affordance in concept composition. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 245–268). London: Springer.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*, 158–190.
- Minda, J. P., & Smith, J. D. (2011). Prototype models of categorization: Basic formulation, prediction, and limitations. In E. M. Pothos & A. J. Willis (Eds.), *Formal approaches in categorization* (pp. 40–64). New York: Cambridge University Press.
- Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science*, *12*, 529–562.
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Nosofsky, R.M. (2011). The generalized context model: An exemplar model of classification. In E. M. Pothos & A.J. Willis, *Formal approaches to categorization* (pp. 18–39). Cambridge: Cambridge University Press.
- Padó, S., & Lapata, M. (2007). Dependency-based construction of semantic space models. *Computational Linguistics*, *33*, 161–199.
- Paivio, A. (1971). *Imagery and verbal processes*. Oxford: Holt, Rinehart & Winston. Retrieved from.
- Paivio, A. (1986). *Mental representations: A dual-coding approach*. Oxford: Oxford University Press.
- Papies, E. K. (2013). Tempting food words activate eating simulations. *Frontiers in Psychology*, *4*, 1–12.
- Pelletier, F. (2017). Compositionality and concepts—A perspective from formal semantics and philosophy of language. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 31–94). Berlin: Springer.
- Pothos, E. M., & Willis, A. J. (2011). *Formal approaches in categorization*. Cambridge: Cambridge University Press.
- Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, *92*, 81–110.
- Ross, B. H. (1984). Reminders and their effects in learning a cognitive skill. *Cognitive Psychology*, *16*, 371–416.
- Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 629–639.
- Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*, 451–486.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*, 214.
- Talmy, L. (1983). How language structures space. In H. Pick & L. Acredelo (Eds.), *Spatial orientation: Theory, research, and application* (pp. 225–282). New York: Plenum Press.
- Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*, 49–100.
- Talmy, L. (2000). *Toward a cognitive semantics, Vol. 1: Concept structuring systems*. Cambridge, MA: The MIT Press.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

- Thill, S., Padó, S., & Ziemke, T. (2014). On the importance of a rich embodiment in the grounding of concepts: Perspectives from embodied cognitive science and computational linguistics. *Topics in Cognitive Science*, 6, 545–558.
- Tomasello, M. (2009). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Werning, M., Hinzen, W., & Machery, E. (2012). *The Oxford handbook of compositionality*. Oxford: Oxford University Press.
- Whitney, P., McKay, T., Kellas, G., & Emerson, W. A. (1985). Semantic activation of noun concepts in context. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 11, 126–135.
- Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, 4, 167–183.
- Wu, L. L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, 132, 173–189.
- Yeh, W., & Barsalou, L. W. (2006). The situated nature of concepts. *The American Journal of Psychology*, 119, 349–384.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Compositionality and Concepts—A Perspective from Formal Semantics and Philosophy of Language

Francis Jeffrey Pelletier

Abstract It's no secret that different of the subfields in cognitive science dispute what the correct solution is to various problems that they each investigate in their separate ways. Sometimes this is due to differing antecedent ideas about what is the appropriate way to investigate the phenomenon, other times it is due to differing antecedent ideas about what principles an adequate solution should embody, and still other times it is due to differing antecedent ideas concerning what the dispute is about... as for example when they use the same terminology for different phenomena. This paper is an investigation into these differing antecedent ideas in the realm of meaning and compositionality as they play out in linguistics, cognitive psychology, and philosophy of language. The focus is on the notions of *subjective meaning* and *objective meaning*, and a main conclusion is that there needs to be a “two-factor semantic theory” to accommodate the overall goals of both sides. Some steps are made towards that end, but various previous attempts are argued to have missed the point. In the end it will be shown that, clearly, more work needs to be done along a number of specified dimensions, especially on the subjective side of the dispute.

1 (Some) Deep Background

I am going to be making some remarks about “what philosophers of language believe” and “what cognitive psychologists believe”—as well as some equally general-sounding claims about “what linguists/formal semanticists/cognitive scientists/etc. believe”. Like all such general statements made about a group, mine should not be taken as universally quantified.¹ (It would be almost beyond belief were all

¹In fact, most researchers into the field of generics think that such “generic statements” can be true—depending on the particulars of the case—even if only a relatively small percentage of the relevant population manifest the property in question, especially when this subgroup is “salient” in some way. See the discussion in Krifka et al. (1995).

F.J. Pelletier (✉)

Philosophy Department, University of Alberta, Edmonton, AB T6G 2E7, Canada
e-mail: francisp@ualberta.ca

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_3

cognitive psychologists to believe the same theoretical things; it would be an even further stretch to think that all philosophers of language believed the same theory; and so on).

I furthermore will be making claims about the notion(s) of compositionality, and will use my explanation of this/these notion(s) as a (rather meagre) web upon which to link what philosophers of language, cognitive psychologists, formal semanticists, cognitive linguists, etc., believe. In fact though, I think that the major reasons for the difference in approaches to these topics come from two other sources, and not compositionality considered simply in itself—which I see as just a consequence of choices within these other sources. The first source is an ontological commitment to one of (what I will call) Atomism and Wholism, and the second source is a semantic commitment to one of (what I will call) Subjectivism and Objectivism. I will try to emphasize these options as the paper develops. At the end, I hope you'll agree that there are numerous “disconnects”, and that on the other hand there are some rather surprising places where agreement is reached.

I start with compositionality.

2 (Some) Features of Compositionality

There are two conceptions of compositionality that make their appearance in the literature. One version, which I'll call *Ontological Compositionality*² describes the way a complex item is put together. The underlying idea is that the complex is “nothing more than” the collection of its parts (and the parts of these parts, etc.), put together in some perhaps new configuration. Underlying this sort of compositionality is the slogan:

Definition 1 (*Ontological Compositionality*) A whole is a compositional system built up only from materials in the parts, with due consideration about how these parts are arranged.

Consider these two theories (don't worry about their correctness as theories—we're only focussing on whether they express ontological compositionality or not).

Suppose assemblage-1 of neurons is active during task-1 and assemblage-2 is active during task-2. Now consider a “supertask” which intuitively involves doing task-1 and task-2 together. Then: is the assemblage of neurons involved in this supertask made up out of, and only of, those neurons involved in either task-1 or task-2? Or does the new supertask bring an entirely new assemblage of neurons to bear on the new task? Many theorists think that, at least after the new supertask becomes a learned, repetitive action, there is a new group of neurons responsible. And thus, motor neural accounts are *not* ontologically compositional.

²Or perhaps a better name would be *Building-Block Compositionality*.

One view of prototypes³ says that a prototype is a structure that has a number of attribute-value pairs. A question of interest in this area is: Given that prototype-1 has a structure [A: value-a; B: value-b; . . .] and that prototype-2 has the structure [C: value-c; D: value-d . . .], is the prototype of the “combined prototypes” (that is, the conceptual combination of the two prototypes) made up only from the material in the two component prototypes, that is, from the attribute-value pairs that are in one or the other of the prototypes? Many theorists answer ‘yes’ to this question, thereby claiming that this manifests the (ontologically) *compositional* view of prototypes. As an example, the prototype of PET FISH would be (ontologically) made up compositionally from the prototypes of PET and FISH because all the “relevant material” (the attributes) could be found in the these smaller, embedded prototypes, and merely needs to be “re-arranged”.

A name for the underlying theory that holders of Ontological Compositionality believe is *Atomism*—the view (in whatever area is under discussion) that there are smallest items (atoms—that is, whatever the area claims to be its minimal elements) and that all other ontological items (in the area) are larger and larger wholes that are ultimately made up of these atoms put together in some specific way. (Different ways will generate different objects out of the same parts.) The denier of Ontological Compositionality believes in *Wholism*—the view (in the area under discussion) that there are some items whose makeup at least partially consists in things that are not in its parts.⁴

Since Wholism takes it that some complex object (of whatever type) can contain things that are not in any of the parts, Atomists naturally find this quite puzzling. One reaction is:

Those who say that a whole is greater than the sum of its parts either don’t know what a whole is, or don’t know what a part is, or don’t know what being a sum is, or don’t know what being greater than consists in.

To take our example of assemblages of neurons, the Atomists would claim that the alleged complex object does *not* have the initial assemblages of neurons as parts. It was a completely different object that was created by learning or practice. It was not the initial assemblages, rearranged, plus something new added on—to make it into a whole. It was just a different object. So it neither obeys nor violates Ontological Compositionality: we just have two independent objects. Or at least, this is the reaction that holders of Ontological Compositionality would claim.

But it is never totally clear when Wholism applies. According to the Boodin quote in footnote 4, any whole that is generated in accordance with gestaltism violates Atomism. Is a heap of sand, made from putting a bunch of grains into close proximity to one another, a *new whole* over and apart from the grains? Or is it just the grains, rearranged by being placed into the relevant proximity? Is a sentence . . . for instance one which is made up of a noun phrase plus a verb phrase . . . such a whole? Did something new come into existence, a *sentence*? Or is the alleged whole actually *nothing more than* the noun phrase and the verb phrase *put together in the appropriate way*? How can we tell in every case whether there is an *entirely new, different*

³E.g., in Hampton (1979, 1982), Smith et al. (1988); the idea can be seen already in Rosch (1975).

⁴The OED cites Boodin (1939) as the originator of the word ‘wholism’: “Two conceptions . . . namely, creative synthesis (or emergence), and wholism (or gestaltism). . .”.

object coming into being, or just the parts that are arranged appropriately? Are the (computational) connectionists right in saying that *every* property of a connectionist system is an emergent property?⁵ We leave these questions aside at this point.

But compositionally of the Ontological variety is not the only notion in the neighbourhood. When discussing linguistic matters, it is far more common, at least in philosophy of language and formal semantics, to see a discussion that presumes there to be an already-given, well-defined system of the sort that Ontological Compositionality has characterized—namely the syntactic structures of language, treated as a complex object made from atoms (and not as a newly-created whole). The atoms are lexical items and the wholes that are made from them are syntactically-defined larger and larger linguistic units. But if we become interested in some *property* of this system—for instance, in the *meanings* of the items—the issue then becomes whether there is a way to define this property—the meaning—of *all* complex items solely in terms of the meanings of the syntactically-given parts of the complex item and their syntactic method of combination. A ‘yes’ answer to this for every complex member signals that the semantic theory assigns meanings compositionally. One might note that in this kind of compositional theory, there are *two* structures: the first, antecedently-given one (in the language case, the syntactic structure), and a second structure (in the language case, it would be the structure that “the meanings” manifest) which is a kind of mirror of the first structure. So another way of asking whether a semantic theory is compositional in this sense is to ask whether there is a homomorphic mapping from the syntactic structure to the semantic structure.

In the language case, this mapping is called “the meaning function”, which most authors writing about semantic compositionality symbolize by μ . So, $X = \mu(A)$ means that (i) A is some element of the first structure [i.e., A is some syntactic item], and (ii) that X is the item in the meaning structure that is paired up with A (i.e., X is the meaning of A). Then this second conception of compositionality asserts that there is a function f such that whenever A is composed (in the sense of Ontological Compositionality) of B, C, D ... by means of syntactic method R, then $\mu(A)$ is f applied to $\langle \mu(B), \mu(C), \mu(D), \dots \mu(R) \rangle$. That is: the system is compositional if and only if, there is a function f such that for every (syntactically) complex item A in the syntactic system, its meaning $\mu(A)$ is a function of, and only of, the meanings of A’s syntactic parts, together with the way they are combined. If NP_1 and VP_1 make up the entirety of S_1 and they are combined by rule-R to do so, then $\mu(S_1) = f(\mu(NP_1), \mu(VP_1), \mu(R))$. Underlying this type of compositionality is the slogan:

Definition 2 (*Functional Compositionality*) The μ of a whole is a function of the μ ’s of its parts and the ways those parts are combined.

⁵I am speaking generically here about connectionists and generically about their views of most properties.

If we restrict our attention to cases where μ yields meanings, then this abstract notion of Functional Compositionality becomes “semantic compositionality”.

A difference between the two notions of compositionality concerns whether some “whole” can contain things not in the parts. According to the ontological building-block view, no; but according to the functional version, yes. For, the first notion allows the whole to contain only what is in the parts, possibly re-arranged in some manner. But the second allows the thing associated with a whole (in the linguistic case: the meaning of a complex whole) to be a *function* of the things associated with the parts (in the linguistic case: a function of the meanings of the syntactic parts and syntactic mode of combination). And then there is nothing to stop such a function from introducing new material into the thing associated with the whole—that is, the function can make the meaning of a whole contain many new and radically different things from what are contained in the meanings of the parts. According to functional compositionality, all that is required is that this be a function—which merely means that it must introduce this same material and generate the same (type of) result every time it is faced with the same (type of) parts and manner of combination.

A simplified example to illustrate the difference between Ontological and Functional compositionality is this. Many theories have thought that adjective-noun combinations were to be analyzed semantically as a conjunction.⁶ Thus, *pink ribbon* is assigned a meaning like PINK(X) & RIBBON(X). Note, however, that we have introduced something into the meaning that is not in the initial phrase nor in the meanings of PINK or RIBBON: the meaning contains &, a conjunctive operator, which is not to be found in the syntactic representation of *pink ribbon*. So, the letter of the definition of Ontological Compositionality would not allow this to happen. On the other side, there is no problem for Functional Compositionality to generate such a representation by saying that the meaning of the syntactic operation ADJPHRASE \Rightarrow ADJECTIVE NOUN is to insert a & (plus add the variables).

However, finding some arbitrary function to do this job is easy—too easy, according to many theorists. For, it seems possible, without any further constraints, to gerrymander the function so as to arbitrarily (but in accord with the function’s definition) admit or omit anything at all as the value of any object in the function’s domain, because a function need not be “regular” or “systematic” or “uniform” or “orderly” in any intuitive sense. For example, Johnson (2014) remarks that while the following three French sentences in (1) have various parts in common, it is still possible to compositionally assign them the respective meanings in (2).

- (1) a. Le chien aboie.
The dog barks.
b. Le chat aboie.
The cat barks.
c. Le chat pue.
The cat stinks.

⁶Although current semantic theory does not hold to this account. I’m using it here merely as an example to illustrate a difference in the two conceptions of compositionality.

- (2) a. The dog barks.
 b. The cat dances.
 c. The skunk eats.

Johnson says that we can see that sharing a verb does not say anything about similarity of meanings, sharing a subject NP says nothing about a similarity of assigned meanings. “Nevertheless,” he remarks, “there exists a function that takes the syntax, and the meanings of the morphemes...and maps it to these meanings....In fact, any random, unsystematic assignment of meanings to sentences is compatible with the functional conception of compositionality....This is ‘dependence’ only in the weakest sense of that word.” In response to this general sort of worry, (Szabó 2012, p. 71) argues for the inclusion of a requirement that the function be one that makes the meanings of the parts *determine* the meaning of the syntactic whole. (Pelletier 2012, pp.160–1) says that the function has to correspond to the way that the meaning of the complexes *really depend* on the meanings of their parts... where ‘really depend’ is taken in some sort of ontologically-explanatory way. Pagin (2009, 2012) suggests that the function has to be a *computable function*, or perhaps an “elementarily computable” function—or even a “humanly-computable-by-anyone” function.

The position that opposes Functional Compositionality is—at least when the realm of meaning is being discussed—(*semantic*) *holism*.⁷ A semantic holist takes it that the meanings of some complex expressions cannot be accounted for by *any* function that “really depends” on the meanings of their parts. We discuss this notion later, in Sect. 6.4.

Let me emphasize that Ontological and Functional Compositionality—as well as their respective contraries, Wholism and Holism—can be applied to many different realms of inquiry. I focussed on the language-meaning case with respect to Functional Compositionality, but that was just for convenience, and because there is a large literature about this application. We return to the topic of compositionality and its various manifestations after introducing a bevy of *dramatis personae*.

3 (Some) Philosophers of Language

It would be very difficult to find any modern philosophers of language who hold that meaning is “in the mind”, or even just the weaker claim that the meaning of a lexical item is a mental concept.⁸ Locke (1690) had said

⁷Note: without the ‘w’. It’s a strange word: it ought to mean a belief in the ontological reality of holes, or perhaps a belief in the virtues of holiness. But it doesn’t. The OED attributes the first occurrence to Smuts (1926).

⁸I *did* say that I’m speaking in generic sentences! There are works such as Hinzen and Sheehan (2013) in which an individual’s concepts are viewed as having been generated pre-linguistically and then having some linguistic label attached to them. But then this label becomes subject to categories/conditions/features of language—both of “universal grammar” and the grammar of the

Besides [the ability to make] articulate sounds, therefore, it was further necessary that [Mankind] should be able to use these sounds as signs of internal conceptions; and to make them stand as marks for the ideas within his own mind, whereby they might be made known to others, and the thoughts of men's minds be conveyed from one to another. *Essay Concerning Human Understanding* iii.1.2

... a word is made arbitrarily the mark of such an idea. The use, then, of words, is to be sensible marks of ideas; and the ideas they stand for are their proper and immediate signification. *Essay* iii.2.1

This is nowadays cited in introductory philosophy of language textbooks as “the most naïve view of meaning”, and the remainder of the text is taken up with a brief rebuttal and then an exploration of alternative theories. Even the few avowedly mentalistic theories (particularly that of Grice 1989) eschew the notion of a mental *concept* in their theory, and employ instead the notion of *intention*, as in “intention to communicate something”, where this “something” is normally something “in or about the world”—such as the proposition that it is raining outside.

Philosophers of language and formal semanticists more generally hold the view that the meaning of a lexical item⁹ is some extra-mental item, which they (unfortunately?) call a concept. Mind you, different ones of these philosophers and semanticists hold different views concerning what such an extra-mental concept is—some hold they are “atomic” and cannot be analyzed. Others think that lexical decomposition is an observable phenomenon, so the corresponding concept must be decompos-

(Footnote 8 continued)

particular system becoming internalized. Likewise there are the various works of Fodor (1975, 1983, 1998, 2000, 2001; among many others) which have advocated a “language of thought” whose atomic members are somehow associated with natural-language lexical items. But the atomicity of the LOT items entails that the linguistic items also must be atomic and not subject to any definition or decomposition. Many theorists have found this implausible. Probably they will find the more recent Fodor and Pylyshyn (2015) even more implausible, where there is *no* intensionality at all in a concept, only extensionality considerations. (That is to say, only membership conditions for instantiating the concept.).

⁹Actually, we mean an individual *sense* of a lexical item. So we really wish to discuss separately, and not as if one word, such cases as: *Pitch: the degree of acuteness or gravity of a tone or sound* at the same time we discuss *Pitch: the central part of a cricket field* or *Pitch: the act or manner of pitching a baseball to a batsman* or *Pitch: a dark-coloured viscous substance* or *Pitch: a stroke with a lofted golf club* and so on and on. Or if that seems obviously to be a group of separate words—and hence obviously separate concepts—to you, consider the case of the noun *Lamb* whose senses include *a young sheep* and *the flesh of a young sheep used as food*. If those still seem as obviously being separate meanings, each to be a separated sense, consider *Collection: the activity of gathering things together* versus *Collection: A group of things gathered together*. The former is a sense indicating an activity; the latter is a sense indicating the result of the activity. Are these separate senses, even though closely related, or not? Some say yes, others say no. If they are separate, then they need to be treated as separate “lexical items”. But all this depends on how finely we would like to individuate senses. That having been said, we will not delve into that topic in this paper. For interesting thoughts on the matter, see Chapt. 8 of Atkins and Rundell (2008). (Both Atkins and Rundell are dictionary editors with many years of experience in issues concerning the individuation of senses.) We also wish only to countenance the “base form” of such a lexical item—a lexeme, as they are sometimes called.

able. Still, the main view in philosophy of language is that the meaning of a lexical item is a concept, in this extra-mental understanding of ‘concept’.

For (Frege 1891, 1892a, b), a concept is a particular type of function: one which, when given a proper name (or as we would say now, a singular term) as an argument, yields a truth-value as a result. In more modern predicate logic terminology, it is a monadic predicate (or rather, that entity in “reality” which corresponds to the predicate—a property, perhaps). Note that this is explicitly *not* a psychological item—Frege is quite resolute in his opposition to “psychologism”.¹⁰

Philosophers of language have antecedently believed that language “talks about the world” in that what one says is to be evaluated for truth or falsity¹¹ against what is happening in the world, not in anyone’s mind. When a person says *X*, s/he is not judged in accordance to how much that person believes that his/her concepts justify the assertion of *X*, but rather we just say the person has spoken falsely, if what they say does not actually match the world. Even if one has a stereotype of SNAKE that includes SLIMY, they cannot truthfully assert *Snakes are slimy* simply on that account. Even if *everybody* had a mistaken belief about some such feature of objects (maybe everyone’s mental idea of SNAKE includes SLIMY), it would not make the claim true. Present tense assertions describe “the way the world *is*”; future tense assertions describe “the way the world *will be*”. There is nothing here about the way the speaker believes it to be, believes it will become, or “conceives” it to be. For, if it were, then we’d say s/he was speaking truly even when saying something false—so long as s/he just believes it. (And while there are doubtless philosophical theories that promulgate such a viewpoint, they don’t include anyone in the philosophy of language.¹²)

This sort of consideration is motivated in part by some famous—and still accepted—doctrines brought out in works by Hillary Putnam (1975) on natural kind terms, by Saul Kripke (1980) on proper names, and by Tyler Burge (1979) on various terms that seem to rely on societal features rather than on an individual’s mental concepts. The moral of these works, and the lesson that has been adopted in philosophy of language, is nicely encapsulated by a phrase made famous by Putnam: “Meaning Ain’t in the Head”. Philosophers of language find it difficult to believe that there are respected (and reflective) theorists who seem to think otherwise.¹³

We can trace the difference between the “objectively oriented” philosophers and the “subjectively oriented” psychologists (and cognitive linguists) that are to be discussed in the next section by means of the *What did s/he say?* test.

Test 1 (What did s/he say?) *When a person utters a sentence, X, ask “What did s/he really say by uttering X?”*

¹⁰For a summary of Frege’s (and Husserl’s) arguments against “psychologism”, see for example Pelletier et al. (2008). These were directed against various attempts to think of mathematics and logic as “psychologistic”, but one might also see them as bolstering his position on thinking that meanings and concepts can’t be mental entities.

¹¹Or along other dimensions corresponding to speech acts other than assertion.

¹²Still talking generically!.

¹³I continue to speak generically!.

If, for example, the speaker utters “All snakes lay eggs”, then when the philosopher that we’ve been discussing applies the test, it would be claimed that the speaker *really said* that every (species of) snake procreates by laying eggs. This is to be evaluated by appeal to worldly facts, and if there are (species of) snakes that give birth to live young, then what the speaker said was false. If you take this view of *what is said*, you will always claim that these cases—where the speaker is sincere, but mistaken—show that meaning is *in the world* and that meaning is not a matter of what the speaker believed nor of what his/her concepts might be contained or was organized in accord with. But since that is what we say about cases when the speaker utters a falsehood, it seems equally to follow that when the speaker talks correctly, so too is his/her meaning a matter of what is in the world.

There are those on the other side, those who think that what the speaker *really said* was a fact about his/her concept of a snake, then the utterance “All snakes lay eggs” might not be false, depending on what features his/her concept SNAKE happened to have. It seems relevant here to note that in ordinary social interactions we distinguish between speaking falsely and lying. The former seems to track what the Objectivist has in mind by the **what did s/he say?** test, and therefore is about “reality”. The latter seems to involve the speaker’s mental state, and depend on whether the speaker is truly reporting what his/her concepts contain (with the intent to deceive). A more stark way of putting this point is that Objectivists take the **what did s/he say?** test to require an answer *about the world* whereas Subjectivists must take it to require an answer *about the speaker’s beliefs*. Furthermore, and although intuitions here can be a bit murky, it seems that if someone intentionally mis-reports his/her conceptual beliefs (say, s/he believes that snakes all lay eggs, but is trying to tell a falsehood) in saying “Some snakes give live birth”, then this is nonetheless a lie even though it is a truth.¹⁴ This intuition would be in accord with what the Subjectivist seems to hold about the **what did s/he say?** test, although I think one should instead identify it with **what did s/he intend?**

4 (Some) Cognitive Psychologists

There have been a number of nice reviews of the recent history of the notion of a concept within cognitive psychology, telling mostly the same story (Komatsu 1992; Laurence and Margolis 1999; Murphy 2002; Machery 2009; Barsalou 2010; from a more recent viewpoint see Hampton and Jönsson 2012; and from a recent philosophical perspective see Leben 2015). The narrative generally goes like this:

¹⁴About 70% of snake species are oviparous (they lay eggs). Of the other 30%, about two-thirds are ovoviviparous (eggs are hatched while within the body then expelled). The remainder are viviparous—they give live birth by nourishing the developing young through a placenta and yolk sac, a very unusual feature for a reptile. While ovoviviparous snakes, like the rattlesnake, are often said to give live birth, this is not true in the same sense that the viviparous snakes do it—which is akin to how mammals give live birth.

In the Classical Version of concepts, a concept was some mental item that classified items into those that fell under the concept versus those that didn't. The general idea was that a concept was a definition, although it was not always clear whether the theorists held that the classifying properties had to indicate *necessary* properties of the item in "the world" (and a sufficient number of them) or whether they might be able to "get along" with accidental properties so long as the classification ability worked in all the cases in which it was called upon. Would it matter whether our concept of orangutan or of baking soda listed exactly the biological classification properties¹⁵ or chemical features (NaHCO_2)—or instead listed a sufficient number of properties that any object that was encountered could be correctly classified as an orangutan or not, or as baking soda or not? [Well, yes it does matter, philosophers say. But that didn't seem to affect any of the psychological theories of the time.]

But data involving typicality judgments and category membership judgments were seen as overthrowing the Classical theory. Even those theorists who wanted to maintain that a concept had a "core" part and a "periphery" (sometimes called a "dictionary versus encyclopedic" distinction) found that there was no special place for such a distinction: it took subjects longer to judge that a chicken was a bird than that a robin was a bird. Presumably this difference is due to "non-core" aspects, such as size and flying ability, since they both are 100% birds, and everyone knows this. Thus there came to be a thought of replacing this with some sort of Prototype representation, one that attempted to accommodate the typicality and membership judgments in terms of "distance from the mental prototype". Here a prototype is something like a "most typical instance" or an "average" of all the instances encountered. But this in turn was refined into (or replaced by) a "family-resemblance" version of the theory. (The reference here to Wittgenstein was intentional, even if these theories were not, on the whole, all that well-grounded in Wittgenstein 1953.)

Another way to accommodate the evidence from typicality judgments was to think of a prototype as an *exemplar*: some particular example of the concept that one happens to fasten on. In addition, there will be a number of learned "rules" about how the exemplar could be modified and still represent a member of the relevant category. When this seemed not to adequately match the full range of typicality judgment data, a more careful version of Prototype theory was elaborated as having feature lists each one with associated probabilities. The concept DOG could have a feature COLOUR with many possible instances listed, each associated with a probability or likelihood. (Note now that there is no special distinction here between "core" and "periphery" of these features.) But computational complexity issues concerning the probabilities of *combinations* of features made these theories be untenable for some,

¹⁵The suggestion in the text of necessary or essential features defining scientific categories is perhaps most appropriate for non-biological concepts, or biological concepts as used by laypeople. But it is not the current viewpoint of professional biologists. Instead the current view is that *taxa*—species, genera, phyla—should be conceived genealogically, in terms of common ancestry with selected paradigms. Thus, for a professional biologist, "being an X" is a relational property of an individual organism, independent of both the superficial properties used for initial identification (amphisbaenias are not snakes despite being legless reptiles) and of explanatory properties (tuna are ray-finned fish despite being warm-blooded). Thanks to Allen Hazen for discussion on this topic.

and gave way to certain variants. If you think that every combination of values among the different feature lists and their probabilities could result in different outcomes for some further feature and its probability, and that this also depends on the probability of an interaction effect between groups of the features, it is pretty easy to see that this is not a tractable problem to characterize and nor to compute a result, if one tries to compute these in a straightforward manner. One way to extend this model could be to add further structure to the concepts, such as having many exemplars and a statement of the particular values each has on each feature dimension. One might then propose that a novel item is evaluated by sampling the stored exemplars and determining how close the newly encountered item is to the value generated from the sampling. But this seems to be no longer a pure prototype theory.

It seemed to many theorists that all these prototype theories have the shortcoming that there was no structure to the feature lists. Schemata theories were therefore devised that divide the various properties envisaged by prototype theories into separate “dimensions” that list the sorts of values they can embody, and can be related to other dimensions or values in those other dimensions. And still more recently “theory based” versions of this same idea have come to the fore with the underlying idea that concepts are not just learned with an already-existing background of other concepts, and therefore have to “fit in” with them, but their internal structure will have to embody previous knowledge about what sort of features are likely to be important in a new concept.

And yet even more recently we’ve seen “neo-empiricist” theories of concepts rejecting the assumption that there has to be *any* relationship among the features of a conceptual structure. Instead we get temporary ad hoc groupings from sensation and motor perception that fit whatever situation one is in. You are in a situation/context where it is important to recall the colour pattern of poisonous snakes, then your ad hoc concept of SNAKE will contain that sort of colour information. But if instead it is a situation where it is important to recall the type of constrictor snakes and their size, then your ad hoc concept of SNAKE will contain that information. The idea is that there is no continuous concept, no such thing as a concept of SNAKE in general. (See Barsalou 2008, 2010, 2017, and in a related framework, Prinz 2002.)

The general idea is that the meaning of some general term (or singular term, for that matter) is the mental concept it occasions. As we have just seen, some of the recent theories hold to the view that such concepts aren’t merely “retrieved” but are actively constructed on the spot as the linguistic and environmental input is received. The construction depending on features of the environment, broadly conceived. We will look at this “construction of meaning” in Sects. 5.2, 6.2 and 6.3 below. In the literature that discusses this—both in the psychological literature and in the cognitive linguistics literature (which is discussed in Sect. 5.2 below)—it is remarkable that focus is given to very simple lexical items, and sometimes to the “conceptual combination” apparent in two-word phrases. But language is “special” (in ways outlined below in Sect. 6.4). In particular it can create arbitrarily long noun phrases that, according to the present theory, should also designate concepts. But one wonders whether the mechanism of conceptual-combination-in-context is up to the job. Here is what Carlson (2010) says about two examples:

[N]ot only do we talk about flying geese and meowing cats and leaf-bearing trees, but we also talk about

- *Unpainted kitchen appliances that are just beginning to rust* need to be replaced within a period of two to three months by a qualified kitchen professional in order to prevent any possibility of bacterial contamination.
- *Friendly but slightly confused medical professionals without appropriate training who nonetheless have medical degrees from top-ranked teaching hospitals* have much to contribute to society beyond their incomplete medical expertise.

The principled point is that linguistic expressions can be unbounded in complexity... with the consequence that if these are mapped directly onto certain types of mental states that we are calling concepts, then we stand in need of a device ... that can produce the arbitrarily large number of corresponding brain states in order to provide these phrases with appropriate denotations.

Depending on how this “device” is defined—and if such a device is really possible to define—it might describe Ontological Compositionality or it might obey Functional Compositionality. Or it might obey neither. But until such a “device” is described, it is not possible to evaluate claims concerning whether the mental combination being presumed by these modern theories of concepts is or isn’t compositional.

5 (Some) Linguistic Theories

Scholz et al. (2016) list three “grand traditions” in modern (and not-so-modern) linguistics, which they label *Essentialism*, *Emergentism*, and *Externalism*. These grand traditions are characterized by their different attitudes towards seven different aspects of what a linguistic theory should be about or achieve. The seven aspects are:

- What are the primary data for theory construction in linguistics?
- What is the primary subject of a linguistic theory?
- What is the aim of a successful linguistic theory?
- What is linguistic structure?
- Besides accuracy, what is to be valued in a linguistic theory?
- What is the relation between child language and adult language?
- What is learned when a first language gets acquired?

According to (Scholz et al. 2016)’s categorization, these grand trends reflect a tendency within different groups of researchers to answer these seven questions in one of three ways. And so the members of a trend can be characterized by commitments to a number of central tendencies or foci. But as Scholz et al. are also at pains to remark, a wide variety of differing research programs could be carried out within each of the three trends... although as a sociological fact, members of each of the three trends seem all to focus on more or less the same directions as one another.

The three trends can be briefly explained by saying that (a) Externalists are most interested in “language as it occurs”, and the most visible research effort of this trend here is corpus analysis, (b) Emergentists are most interested in how language is used by people and how this is an outgrowth of other cognitive aspects of people, and probably the most visible research effort of this trend is Cognitive Linguistics/Semantics (but see footnote 17 below), and (c) Essentialists are most concerned to presume an internalized language faculty that can be characterized by an abstract, computationally-oriented model. The two most visible subgroups of this grand trend are the theories put forth by Chomsky over the decades and the formal semanticists. The three groups’ answers to the seven aspects of what a linguistic theory should be like are given in (Scholz et al. 2016)’s Table 1.

5.1 (Some) Externalists

As Scholz et al. (2016) characterize these three trends or tendencies of research goals, the Externalists are most typically exemplified by some computational corpus linguists, whose goals are to find underlying patterns that are discernible in actual usage when the sample is large enough to show it. The patterns might then be characterizable by some rules that could be used to give an overall theory of the language. Another version of Externalism holds to certain philosophical accounts of meaning in terms of use: for example the “inferentialism” of Brandom (1994, 2000) or the “language game” and “form of life” of Wittgenstein (1953). Many forms of linguistic behaviourism could also be characterized in this way, for example Skinner (1957) and in a different way Osgood et al. (1957). The computational approaches to linguistic theorizing in the Externalist vein tend to eschew the postulation of mental items, such as concepts, intentions, purposes, and the like, and instead take meaning to be the likelihood of co-occurrence with nearby words. As can be seen from the ‘AIM’ row of Table 1, it is not a central concern of Externalism to worry about generating semantic representations that obey semantic compositionality, nor with psychological data concerning speaker/hearer mental life during conversation. And hence they stand outside the sort of discussion this paper is concerned with, namely the interaction between linguistic features and mental items. So, I will ignore the work in the Externalist direction in the remainder of this essay, except for mentioning how some inferentialists view semantics.¹⁶

¹⁶Important references for current versions of Externalism are (Manning and Schütze 1999; Jurafsky and Martin 2008) and (Schubert 2015, esp. Sect. 9: “Statistical Natural Language Processing”). In this section, Schubert discusses how this is a new direction in computational linguistics that “is significant from a philosophical perspective, and not just a practical one.” Schubert discusses a wide range of areas where this ‘philosophical shift’ has been applied.

Table 1 Three approaches to the study of natural language. Printed with the kind permission of the Editor of *The Stanford Encyclopedia of Philosophy* and the authors of Scholz et al. (2016)

	Externalists	Emergentists	Essentialists
Primary phenomena:	Actual utterances as produced by language users	Facts of social cognition, interaction, and communication	Intuitions of grammaticality and literal meaning
Primary subject matter:	Language use; structural properties of expressions and languages	Linguistic communication, cognition, variation, and change	Abstract universal principles that explain the properties of specific languages
Aim:	To describe attested expression structure and interrelations, and predicting properties of unattested expressions	To explain structural properties of languages in terms of general cognitive mechanisms and communicative functions	To articulate universal principles and provide explanations for deep and cross-linguistically constant linguistic properties
Linguistic structure is:	A system of patterns, inferable from generally accessible, objective features of the use of language	A system of constructions that range from fixed idiomatic phrases to highly abstract productive types	A system of abstract conditions that may not be evident from the experience of typical language users
Value especially:	Accurate modeling of linguistic form that accords with empirical data and permits prediction concerning unconsidered cases	Cognitive, cultural, historical, and evolutionary explanations of phenomena found in linguistic communication systems	Highly abstract, covering-law explanations for properties of language as inferred from linguistic intuitions
Young children's language is:	A nascent form of language, very different from adult linguistic competence	A series of stages in an ontogenetic process of developing adult communicative competence	Very similar to adult linguistic competence though obscured by cognitive, articulatory, and lexical limits
What is acquired:	A grasp of the distributional properties of the constituents of expressions of a language	A mainly conventional and culturally transmitted system for linguistic communication	An internalized generative device that characterizes an infinite set of expressions

5.2 (Some) Emergentists

It is the linguists from the Emergentist group—so-called because they see language as developing out of and emerging from antecedently existing cognitive, social, or biological forerunners—who are the most self-conscious employers in linguistics of

“mental terminology” such as (mental) *concept*. One group—the self-named *cognitive linguists*—is explicit, indeed insistent, about their orientation, that psychological evidence is part and parcel of their own enterprise¹⁷:

...[W]hat is the precise meaning of *cognitive* in *Cognitive Linguistics*, and how does this meaning differ from the way in which other forms of linguistics conceive of themselves as being a cognitive discipline? ... Cognitive Linguistics is the study of language in its cognitive function, where *cognitive* refers to the crucial role of intermediate informational structures in our encounters with the world. Cognitive Linguistics is cognitive in the same way that cognitive psychology is: by assuming that our interaction with the world is mediated through informational structures in the mind. It is more specific than cognitive psychology, however, by focusing on natural language as a means for organizing, processing, and conveying that information. (Gaeraerts and Cuyckens 2007, pp. 4–5)

Sometimes, though, the use of different names by various authors within the framework for the mental entities that they assume to be the mental reality that language is describing or reflecting, can make a casual reader think that they may not be talking of the same thing that cognitive psychologists are. For instance, their use of *schema, constructions, spaces, categories, frames, prototypes, figure/ground*, among others, in explaining the inner workings of their theory may hide this. But about them, under whatever name, the thought is that the meaning or semantic value of lexical items (especially) are these mental items. Some programmatic but representative remarks about the standpoint of Cognitive Linguistics, as made by some of its practitioners, are:

The experience that users have with language shapes cognitive representations, which are built up through the application of general principles of human cognition to linguistic input. The structure that appears to underlie language use reflects the operation of these principles as they shape how individual speakers and hearers represent form and meaning and adapt these forms and meanings as they speak. (Bybee and McClelland 2005, p. 382)

Representation is perhaps *the* most important, and most contested, foundational concept in modern cognitive science. Cognitive Linguistics takes the view that linguistic structure is motivated by conceptual representation and communicative function, thereby placing the representational function of language at the center of its concerns. (Sinha 2007, p. 1280)

[The psychologist’s notion of] the basic level of categorization... is linguistically basic because basic-level terms tend to be morphologically simpler, to be acquired earlier by children, to be used as the unmarked choice for introducing referents into discourse, and to provide the raw material for extensions of the lexicon by means of metaphor, metonymy, and word formation. (Schmid 2007, p. 124)

Several general cognitive abilities can be brought under the broad title of “construal operations”. At issue is the fact that linguistic expressions do not, and cannot, designate a state of

¹⁷There are other schools or orientations within the broadly Emergentist category, such as Functionalism (Bybee 1998; Dirven and Fried 1987), Construction Grammar (Croft 2001; Goldberg 1995; Fried and Östman 2004; Östman and Fried 2004; Kay 1995), various sorts of Sociolinguistics (Halliday 1978; Hurford 2000; Labov 1966, 1975; von Savigny 1988; Tomasello 1998; Van Valin 2003), and Language Evolution (Burling 2005; Cheney and Seyfarth 2005; Greenhill et al. 2010; Li 2002; MacWhinney 2005; O’Grady 2008; Power 1998; Wildgen 2008; Worden 1998).

affairs as it “objectively” is; rather, the scene must be processed and conceptualized by the human mind. Construal operations include attentional processes, force-dynamic construals, metaphor, image schemas, and conceptual blending. (Taylor 2007, p. 574)

The semantic views held by Emergentists can perhaps be gleaned from the foregoing quotations, but in short, and at least when restricted to a lexical item, the view about meaning is:

Definition 3 (*Linguistic Emergentism*) The meaning of a lexical item is a (mental) concept.

Of course, Emergentists will have interesting and complex things to say about the nature of these mental concepts, and their mutability in context—where by ‘context’ is meant not only the longer sentence or discourse in which the lexical item occurs but also the situation when and where it was uttered, and also by the sum of background experiences the speaker has plus his/her beliefs concerning the background experiences of his/her audience. It is one of the fundamental views of Emergentism, or at least the Cognitive Linguistics subtype of Emergentism, that linguistic expressions are widely variable in their meanings, depending on the sorts of factors just mentioned. But then since they are also committed to the view that meanings are mental items, it follows that these concepts must also be able to vary along the same dimensions.

Evans (2009) states a general feature (‘literalism’) he perceives in “traditional linguistic theories”,¹⁸ and which he is concerned to overthrow. It can be seen that Evans’ view of compositionality, which he sees as inherent in these traditional theories, amounts to what I called Ontological Compositionality with its “monotonic”, “additive” view of composition, and not the Functional Compositionality that I had attributed to these other types of linguistic trends. In this I think Evans is wrong, and I think that the notion embraced by formal semanticists (who I shall introduce shortly) is Functional Compositionality. I also see this as one of the “disconnects” between cognitive linguistics/psychology and philosophy of language/formal semantics. A further feature of Cognitive Linguistics, emphasized here by Evans, is that there is no such thing as “the meaning of a word” (or other linguistic item): they are always infected with contextual aspects when they are used, and furthermore when they were first learned. A meaning, on this view, is something fleeting—it comes and then goes as the context (linguistic or situational) changes.¹⁹

[T]he view of word meaning adopted under literalism is that word meanings are assumed to be relatively fixed and stable. Put another way, the semantic primitives which make up a given word meaning can be identified independently of context. Once identified, word meanings are integrated, by applying the rules of the grammar, in order to provide sentence meaning. Literalism, then, assumes that the contribution of language to meaning construction is essentially additive in nature, positing grammatical principles which ensure that the

¹⁸By which he probably intends what I’ve called Essentialism.

¹⁹Actually, Evans tries to separate “senses stored in semantic memory” from those “meanings constructed ‘online’”. But the idea is that the former serve only to guide the construction of the latter, which is the meaning of the word-in-context.

semantic units which result are unable to change or delete the meanings of the units which are conjoined to form a larger semantic unit or expression. This restriction serves to make a larger expression, for instance a sentence, monotonic with respect to its component parts, where the term “monotonic” has to do with the view that the component parts retain their original meanings in the larger expression. Thus, the individual word meanings do not alter their meaning in the larger semantic units of which they form part. (Evans 2009, p. 6)

Later in the book, Evans returns to this thought:

[W]ords, I suggest, do not mean independently of context. Thus, the fundamental problem with literalism is that it attempts to artificially divorce (word) meaning from (situated meaning in) context of use. More precisely, literalism lives in something of a fool’s paradise. It holds that language users retain an idealized, timeless meaning for *open* which they neatly keep apart from the situated meanings of *open* which arise from its use in examples . . . The mistake that literalism makes, then, is in being reductionist and simplistic about meaning.

[W]ords are never meaningful independent of the utterance in which they are embedded, and the encyclopaedic knowledge and extra-linguistic context which guide how words embedded in an utterance should be interpreted. Indeed, evidence from the perspectives of social psychology, cognitive psychology, interactional sociolinguistics, cognitive linguistics, corpus linguistics, and computational linguistics reveals that the view that words constitute fixed, context-independent structures, and that meaning construction is appropriately modelled in terms of the straightforward approach to compositionality sketched above is untenable. (Evans 2009, pp. 21–22)

This viewpoint is not idiosyncratic to Evans. Rather, as he says, part of this general viewpoint of Cognitive Linguistics is that mental concepts, and hence the linguistic items that receive their meaning from these concepts, are determined not by some “core” dictionary meaning but rather are related to much more about the world of one’s acquaintance. Using other terminology, a concept is said to embody *encyclopedic knowledge* and to be augmented or restricted by situational features of the utterance. Many of the early Cognitive Linguists have also promulgated this view, but Langacker (1987; 2008) is perhaps the most theoretically-oriented of the early practitioners, and he was probably the most influential developer of “encyclopedic semantics”. The idea is that lexical items (and longer linguistic phrases) directly access concepts, and since these concepts are developed through interaction with the world in non-uniform ways, these linguistic items can’t be understood independently of the encyclopedic information held in these concepts. Semantic structure, then, is just *identical* to conceptual structure.

Evans is concerned throughout his works to view meaning not only as mental concepts, *à la* Langacker, but to show that these concepts are continually changing, as the following quotation from a later work illustrates:

Language works extremely well because it is part of a larger meaning-making complex. It’s dependent upon a suite of other capacities – in particular, a repository of thoughts and ideas, that we carry with us, in our minds – upon which language draws each time we open our mouths to speak. Language, together with this mental apparatus, co-conspires to produce meaning, enabling effective communication. (Evans 2015, p. 6)

(One notes that, even were all this true about how the co-conspiracy works, it is quite unclear how that will help with communication!)

This group of the Emergentists—the Cognitive Linguists—is one of the two groups of linguists that I wish to keep in focus during the remainder of this paper. We will return to a discussion of this group, with an eye to seeing whether and how it is similar to (some) work on concepts in cognitive psychology of the sort I described above. But first we continue our account of linguistic theoretical frameworks.

5.3 (Some) Essentialists

A large number of contemporary Essentialists who follow Chomsky's teaching on this matter claim that semantics and pragmatics are not a part of the study of language.

[T]he study of meaning and reference and of the use of language should be excluded from the field of linguistics ... [G]iven a linguistic theory, the concepts of grammar are constructed (so it seems) on the basis of primitive notions that are not semantic (where the grammar contains the phonology and syntax), but that the linguistic theory itself must be chosen so as to provide the best possible explanation of semantic phenomena, as well as others. (Chomsky 1977, p. 139)

It seems that other cognitive systems – in particular, our system of beliefs concerning things in the world and their behavior – play an essential part in our judgments of meaning and reference, in an extremely intricate manner, and it is not at all clear that much will remain if we try to separate the purely linguistic components of what in informal usage or even in technical discussion we call 'the meaning of linguistic expression.' (Chomsky 1977, p. 142)

In Chomsky's view, "it is possible that natural language has only syntax and pragmatics" (Chomsky 1995, p. 26); that is, only "internalist computations and performance systems that access them"; semantic theories are merely "part of an interface level" or "a form of syntax" (Chomsky 1992, p. 223). In his view, then, there is no distinction between semantics and pragmatics, and *neither* forms a part of linguistic theory.

So, in the Chomskian take on the matter, semantic topics such as the meaning of lexical items are a matter for theories other than linguistics to determine. It is a matter of the cognitive systems that describe our beliefs about things in the world which would determine what meaning (and reference) is. Without using the term *concept* in any technical sense, it would seem that Chomsky endorses the claim made in Definition (3) above, about Emergentism, and denies only that this should have anything to do with linguistics. (Of course, when an Emergentist hears this they will be astonished to hear the 'only' in such a sentence, since they believe that this is the fundamental starting point for linguistic theory!) Nonetheless, we can see that this form of Evans' "traditional linguistic theories" does *not* embody his "fool's paradise of literalism". Meanings and such are not at all a part of linguistic theory.

However, not every Essentialist agrees with Chomsky on this point. Many believe that every theory *should* incorporate a linguistic component—semantics—that yields meanings, in much the same way that most philosophers of language believe there to

be such a separate component of a theory of language. Often, although not always, this component amounts to a truth-theoretic account of the values of syntactically-characterized sentences. This typically involves a translation of the natural language sentence into some representation that is “intermediate” between natural language and a truth-theory—perhaps an augmented version of first-order logic, or perhaps a higher-order intensional language. And other times it involves a further level of syntax, LF (sometimes this is inaccurately called “logical form”), which itself is then interpreted by some truth- or information-theoretic method. The Essentialists who study semantics in such ways usually agree with Chomsky in seeing little role for pragmatics within linguistic theory. But their separation of semantics from pragmatics allows them to accord semantics a legitimacy within linguistics itself, and not “just” in psychology or sociology. This group are known by the general name of Formal Semanticists. And perhaps this is Evans’ group of “traditional linguists” who adopt “literalism”?

So for these Formal Semanticists, such features as the individual differences between speakers on their beliefs about “the world” will become a part of pragmatics, and not a part of semantics. And hence, *not* a part of linguistics. (But rather, of some social psychological theory, or perhaps a theory about an individual’s cognitive life.) This is a view they share with the philosophers of language who were surveyed earlier. Chierchia and McConnell-Ginet (2000) and Heim and Kratzer (1998) are leading textbooks in this general framework, and both adopt a “logic-oriented” approach to semantics that employs a possible world semantics and other formal tools. Their differences reside in their length (Chierchia and McConnell-Ginet is much longer than Heim and Kratzer) and to some extent in their differing underlying syntactic frameworks. (Chierchia and McConnell-Ginet 2000) put their case in the preface and beginning of the book as follows:

[W]e focus on what has come to be known as logical, truth-conditional, or model-theoretic semantics. This general approach to meaning was developed originally within the tradition of logic and the philosophy of language and over the last twenty years or so has been applied systematically to the study of meaning in natural languages, due especially to the work of Richard Montague. (Preface)

Whatever linguistic meaning is like, there must be some sort of compositional account of the interpretation of complex expressions as composed or constructed from the interpretations of their parts and thus ultimately from the interpretations of the (finitely many) simple expressions contained in them and of the syntactic structures in which they occur. . . . An important test of a semantic theory is set by compositionality. Can the theory generate the required interpretations for complex expressions from a specification of interpretations for the basic items? As we will see, explicit specification of how word meanings are combined to produce sentential meanings is not a trivial task. (pp. 6–7)

Somewhat later in the book, this general train of thought is revisited, arguing further that formal semantics is committed to Objectivism and is opposed to Subjectivism:

Meaning manifests itself in the systematic link between linguistic forms and things that we speak of or talk about. . . . Were languages not to provide for significance in this sense, the question of meaning would hardly arise.

Language enables us to talk about the world, to convey information to one another about ourselves and our surroundings in a reliable fashion. What properties of language and its uses underlie this remarkable fact? What allows language to serve as a guide to the world and to enable us to learn from what others have perceived (seen, heard, felt, smelled) without having to duplicate their perceptual experience ourselves? Informational significance does not require that language links to the world in ways that are predetermined by the physical structure of our environment. Nor does it require that environmental information is simply registered or received without active input from perceiving and thinking human minds. Yet it does probably require a regular and systematic correspondence between language and the shared environment, what is publicly accessible to many different human minds. If you are skeptical about informational significance, consider the use of language in giving directions, warnings, recipes, planning joint activities, describing events. Things occasionally misfire, but by and large such uses of language are remarkably effective. Language could not work at all in such ways were it not imbued with some kind of informational significance, being about matters in a public world. (pp. 11–12)

It is also appropriate here to mention the area of “Truth-Functional Pragmatics” (Recanati 2012). This is also an approach that I would like to bring under the wing of Formal Semantics, even though the practitioners think of themselves as allowing “pragmatic” effects to “intrude” on the semantic understanding of utterances. François Recanati is possibly the most influential member of this group, and although this approach seems to endorse non-compositionality (in the Functional sense), Recanati himself has given some ways to have it be “almost compositional” (in Recanati 2012; see also Pelletier 2013 for further ways). On the dimension of Cognitive-Linguistics-to-Formal-Semantics, truth-functional pragmatics is quite close to Formal Semantics, and I will include them into the Formal Semantics group, even though it also shares a number of features with Cognitive Linguistics (e.g., the importance of context in determining the *content* of what is said, and a penchant for claiming that this entails that the “true meaning” of what is said is full of “pragmatic enrichment” and “modulation”). For, the fact is that in these theories the final evaluation of an utterance depends on truth: that is, upon whether the meaning of what is said is to be evaluated against what is happening “in the world”. And this puts it pretty squarely in the camp of the Formal Semanticists.²⁰

This Formal Semantics part of the Essentialist camp is the second group of linguists that I wish to engage about the notion of compositionality, comparing their view(s) with those from philosophers of language, cognitive linguists (Emergentists), and cognitive psychologists.

²⁰Of course, there are differences within the group and thus between (some of) them and (some of) the earlier groups we distinguished. For the purposes of this paper, though, I will focus on their mutual endorsement of Atomism and through that commitment, their advocacy of semantic compositionality and its relatives.

6 (Some) Objectivists Versus (Some) Subjectivists

There is a similarity, as I have noted, between the philosophers of language and the formal semanticists. They both think of “meaning” as being *outside* of speakers, as *public*, as being *in common*. Despite whatever other differences there may be between and amongst them, I therefore will call this merged group *The Objectivists*.²¹

There is a different similarity that I noted between the cognitive psychologists and the cognitive linguists. They both think of “meaning” as something that is *inside* speakers, *private*, as being *personal*. Despite whatever other differences there may be between and amongst them, I therefore will call this merged group *The Subjectivists*.

Carlson (2010) ends his discussion of the difference in usage of the term ‘concept’ in the formal semantics literature and the experimental psychology literature with

Concepts, as least as studied by psychologists, have not found a natural place in the study of natural language semantics in the formal semantics tradition as practiced in the past thirty or so years. This does not point to any shortcomings in semantic theory, I emphasize, for there are excellent intellectual as well as practical reasons as to why this should be. However, it does leave us with something of a disconnect between what is going on in the mind – and I take concepts to be major features of it – and the way we talk about things. (p. 33)

I remarked earlier that philosophers of language—who also employ the word ‘concept’—think of it as being “in the world” and not “in the mind”. Another way of putting all this is to say that Objectivists think of a concept as something that can literally be apprehended or grasped by more than one person (and not as some sort of “abstraction from, or averaging of, the private possessions of many different people”). On the other hand, Subjectivists are committed to a view of concepts that are “private”, in the sense that it is literally true that a concept is in the mind of exactly one person. (Correspondingly, the sense in which different speakers have the same concept is one of “determining similarities between and among the different private concepts, either by abstraction or averaging”.)

This difference is one of the root causes for the disputes that dot the literature. It gives rise to one of the standard arguments used by Objectivists (mostly the philosophers of language) against the Subjectivists (Sect. 6.1). It also gives rise to one of the standard arguments used by Subjectivists (mostly the cognitive linguistics) against the Objectivists (Sect. 6.2). In Sect. 6.3 I point to another place where (some) Subjectivists and Objectivists battle: general terms. I argue in favour of the Objectivists, although I also point to a possible two-tier theory imagined in Sect. 7.3. In Sect. 6.4 I discuss the so-called “Master Argument(s)” in favour of semantic compositionality, bring out a feature that seems to have gone unnoticed by its proponents—that, as an argument from the Objectivist side, it undermines itself by its own presuppositions. I also point out that these Master Arguments do not really support semantic compositionality, but rather they are arguments for some sort of semantic Atomism. But semantic Atomism is compatible with some Subjectivist frameworks.

²¹Yes, there are various other groups that are also called ‘Objectivists’, but I’ll not be talking about any of them in this essay.

We start with a standard argument used by Objectivists against Subjectivists, to the effect that the Subjectivist notion of a concept implies an untenable position concerning “privacy”. If this argument is right it puts the Subjectivists in a very dire position indeed—they would be committed to the incoherence of all of what is believed about human interaction. Let’s see how that goes:

6.1 Concepts and Privacy

Concepts, on the Subjectivist account, are private—there is no literal sharing, no joint ownership, etc. There is thus no direct way to determine whether a speaker and his/her audience share anything whatsoever when it comes to a concept—even in the case where the concepts are occasioned by the same verbal signal (word). And even though one might think of behavioural correlates such as classifying the same objects and seeing how many they disagree with, this doesn’t give any reasonable guarantee of sameness of concept, as opposed to agreement in behaviour.

But perhaps the Subjectivist would say that agreement in (certain types of) behaviour is sufficient for our purposes here; that we don’t really need shared concepts, only shared behaviour of the relevant sort. To this an Objectivist counters by asking whether we really would like to claim literal identity of different people’s sensation of pain, whenever they agree to rate the same number out of 10 on a Numeric Rating Scale. The Objectivist thinks that no one, not even the most committed Subjectivist would agree to that. Indeed, once the Classical theory of concepts is thrown out, it is a part of the Subjectivist story that the concepts are developed by accommodating to the experiences that an individual has. Surely this guarantees that the concept that one person has will be both numerically and qualitatively different from the next person’s. And if we are supposing, as the current theorists do, that the meaning of a lexical item is the concept it occasions, then of course no two people will have the same meaning for any word. Subjectivists have not talked much about whether there are also different ways for concepts to combine in different people on account of their differing experiences, seemingly treating any combination as if it were a conjunction. But even if they were able to give us an account of how simple concepts can combine in whatever ways they do, and how the resulting complex concepts can themselves combine further, and yet further, until a “complete thought” (corresponding to a full declarative sentence) is formed—as I say, even if that is all granted—then since these combining rules or abilities will themselves be due to experience that is learned-by-an-individual, they will thus almost certainly differ from one person to another. In short: this aspect of the Subjectivist theory guarantees that there will be no understanding between any two people, at least if understanding amounts to having (qualitatively) the same concept.

It is tempting to say, and many Subjectivists have,²² that communication does not require “absolute identity” of the corresponding mental items but only that they be “similar enough”. But this can’t really be made out in a non-circular way, it seems to me. The hypothesis that two different minds are “similar enough” with respect to their understanding of some term—say, ‘democracy’—has no empirical content other than simply some antecedent belief that the two people *are* understanding what each other says when they talk, despite their differences in acquisition of the relevant concepts. But that was precisely the (alleged) fact that the “similarity” was hypothesized to explain. The very nature of the Subjectivist notion of concept makes it impossible to give an *independent* criterion of “similar enough”. Contrast this with the Objectivist claim that there literally is something—something *in reality*—that the participants both grasp. It may be correct to object that this is asserted without any independent evidence, but at least it is not a restatement of the issue, as I claimed (in the previous footnote) that the Subjectivist is committed to.

6.2 Concepts and Meaning

Although Objectivists seem to believe they have knock-down reasons for holding that “meaning just ain’t in the head”, as we outlined in Sect. 3, it seems equally obvious to Subjectivists that without taking into account the mental life of speakers (and possibly hearers, too) it is impossible to give an account of what meaning is. After all, isn’t it obvious that a speaker has some belief that s/he wishes to get across, or warning that s/he is trying to convey, or emotion that s/he is expressing, or And are these not all mental states? How could there be any meaning in an instance of language use that did not give any role whatsoever to these mental items? Isn’t it obvious that when a speaker says that *The object in the corner is a chair* that they

²²Here are two, chosen from some 200 years apart:

Men do not understand one another ... by mutually occasioning one another to produce exactly and completely the same concept; they do it by touching in one another the same link in the chain of their sensory ideas and internal conceptualizations, by striking the same note on their mental instrument, whereupon matching but not identical concepts are engendered in each. (von Humboldt 2000, p. 152)

Since the cumulative experience that makes up the conceptual system will differ between individuals, often considerably; so, the ability of two individuals to communicate effectively depends on there being a minimum degree of overlap between their ongoing representations and, therefore, the concepts from which these representations derive. The question then becomes: How much overlap is enough. Critically, understanding one another in communication does not require identical representations between speaker and listener. ... just representations that overlap enough to achieve the current communicative goal. (Connell and Lynott 2014, p. 400)

I think that anyone can see that neither account gives anything more than a restatement of the actual issue, not a solution.

are identifying some mental representation (a percept?) of some object taken to be in the corner, and are classifying it as falling under the concept CHAIR? Given all this, it certainly seems obvious that there is some relationship or other between concepts and meaning. Objectivism denies this, and is deficient for this reason, the Subjectivist says.

The intuitive and natural manner to characterize meaning in a Subjectivist framework is to say that the basic way for mental concepts to operate is by *representing the world*—at least the world as seen from the point of view of the holder or owner of those concepts. A linguistic item then will designate or mean or be brought forth by (or will bring forth) the relevant mental concept of the language user, and will thereby *represent* the relevant aspect of the world. This is a type of “two-level” or “dual aspect” or “two-tier” semantics, and we will canvass that sort of response in the section below (Sect. 7). For now we just remark that the sort of two-tier theory being alluded to also has analogies with theories mentioned in McNally and Boleda (2017) and Winter (2017), and perhaps shows a remarkable convergence of opinion among theorists with widely different starting points.

One formal approach to subjectivist theories of meaning is exemplified by so-called proof-theoretic semantics. Such theories have started by thinking of the meanings of the logical connectives as being given by the ways they can be introduced into a discourse and what sorts of “language moves” they can justify. For some technical reasons this led to adopting intuitionistic logic as the underlying framework for a theory of the use of such connectives (see especially Dummett 1991). Fuller versions of this idea try to generalize the range of such introduction and exploitation moves that can be given a formal exposition. Most of these theories have their formal roots in the intuitionistic type-theory of Martin-Löf (1980) and Prawitz (2006); see for example its development in natural language structures by Francez and Dyckhoff (2010); Francez et al. (2010); Francez and Ben-Avi (2015); Francez (2014, 2015). In this approach, meaning is taken to be a certain type of proof, namely a set of canonical derivability conditions—stated in natural deduction format from which the derivability conditions are based. The idea is that the proof system reflects the “use” of the sentences in the linguistic fragment under consideration and thus allows recovery of their entailment and assertability conditions. This formal conception of semantics is opposed to the more usual formal version in linguistics and philosophy of language of describing the meanings as truth-conditions in arbitrary models, which are taken to be ways the world actually *is*, or might be (the latter for intensional language). In proof theoretic theories there is no notion of “truth” or of “modelling the world”. There are only rules for the appropriate *use* of language. In more informal versions of proof theoretic semantics (Brandom 1994, 2000) this general strategy is called *inferentialism* and is seen as a “use-based theory of meaning”, which employs notions such as “language entry and exit rules”, and “language evaluation rules”.²³ (A nice description of both the formal and informal aspects of the topic is in

²³“Use theories of meaning” trace their ancestry to ordinary language philosophy, especially the sort that is usually attributed to the later Wittgenstein (but now moved to a higher level of formality). In the Wittgensteinian version—and generally carried into the modern versions—the issue of

Schroeder-Heister (2016). We shall not here evaluate the various arguments from the proof-theoretic and model-theoretic sides that allege to find shortcomings in the other approach. A lot of this back-and-forth mirrors the Objectivist–Subjectivist argumentation we have been canvassing, except done on a highly formal and mathematical level. Instead I will simply appeal to proof-theoretic semantics below as a possible formal account of the ways in which people can *use* the representations afforded by concepts to produce an “internal meaning” in our discussion of two-tier semantic theories of Sect. 7.3.)

6.3 *Specificity and Meaning in Particular Instances of Language Use*

Another feature of the dynamic between Objectivists and Subjectivists that holds back some sort of rapprochement is their differing underlying notions concerning just what meaning should *be* in particular cases. We’ve already seen the large-scale differences in the form of context-dependent versus independent, encyclopedic versus dictionary, and the like. But even in a specific case of meaning-in-a-very-definite-context there are important differences between the two camps. Here I outline the Objectivist side against the Subjectivist. The differences come in two forms.

6.3.1 *Generality*

Objectivists and Subjectivists differ on the “level of specificity” that should be a meaning in the various cases. It is not so clear to me why this should follow from their overall positions, but their writings leave little room for doubt that it is so. Objectivists have a notion that some meanings include others as special cases or as particular ways that the meaning can be manifested. For example, they might say that ‘The cover of this book is red’ can be made true in many different ways: it could be crimson, it could be cardinal red, it could be scarlet, it could be cerise, etc., etc. And, they hold, a person could truly say and mean that the book cover was red even all the while *knowing* (or believing) that it has some specific shade, say raspberry. Their Objectivist view is that *what was said* allows for instantiation in many different ways; the speaker’s particular knowledge of the shade is irrelevant to *what was said* (as is any potential hearer’s knowledge of the situation which might dictate the particular shade). Similarly, should some speaker claim ‘There is a red fruit in my refrigerator’, it wouldn’t matter to *what was said* that it was, or was known by the speaker to be, an apple with a red-delicious coloured exterior, a watermelon with a ruby-coloured interior, or a burgundy-coloured grape, etc., etc. The idea, again, is that language

(Footnote 23 continued)

Objective versus Subjective was said to rest on a common mistake, that of not according primacy to how (ordinary) people (ordinarily) use language.

contains *generality*, allowing us to talk without having to always and completely describe each and every detail of what is being asserted.

Presumably this attitude follows from the Objectivist's viewpoint that *what is said* (think of the **what is said?** Test (1) of Sect. 3) determines a certain group of possible states of the world—those possible states that would render the statement true (for assertions; *mutatis mutandis* for other speech acts). So long as the actual world is one of those possible states, the statement is true. Thus, for the Objectivists, the meaning is the set of possible ways that the sentence could be true. And with 'The cover of this book is red', any of hundreds of specific shades would be adequate.²⁴ This picture is nicely captured by Sainsbury (2001):

'Italian *F*' is satisfied by a satisfier of *F* which is of or pertaining to Italy. This is an unspecific but definite, unambiguous and complete meaning. An Italian book is one of or pertaining to Italy, and a book may pertain to Italy by being on my pile of books to take there, or by being about Italy, or by being manufactured in Italy, or by being written in the dominant language of Italy. . . . 'Feline care' is satisfied, in the same sense though not in the same way, by a vet tending a cat and by a cat tending the puppy to which she is acting as foster mother. Both vet and cat supply care of, or pertaining to, felines . . .

Sainsbury's main point in this is the contention that theorists should distinguish between different 'readings' or meanings of a sentence from different ways in which it could be made true. Just because (for example) *This is an Italian book* can be made true by this book's being about Italy, or written in Italian, or the book bought while visiting Italy, and so forth, it does not at all follow that this shows that there are somehow different 'meanings' involved with the sentence. And if these are all the same 'meaning', then a compositional semantics needs only to generate a meaning for *Italian book* that is a function of the meanings of *Italian* and *book*. It is not the job of semantic theory to generate all the different 'readings' of a sentence.

Some Subjectivists (especially in Cognitive Linguistics or where they follow the general (Wittgensteinian) doctrine that all experience is "of the particular" and that a desire for generality is "the root cause of all philosophical conundrums") seem to refuse to acknowledge that there is any reasonable notion of generality in meaning—whether it be in language or in concepts. The fact that someone maybe has in mind that they are attracted to the colour red in their footwear makes it be that when the person says "I'm going to go buy a pair of shoes this afternoon" s/he *means* that they are going to buy a pair of *red* shoes. The fact that some particular physical phenomenon can be described in differing ways is claimed to give rise to different "construals" (this is a technical term).²⁵ Such different construals are seen by these

²⁴A caveat should be registered here, concerning homography and homophony, where different words are spelled or pronounced the same. Different words can give rise to different meanings and hence divergent ways they could be made true. But the same word-sense just gives rise to one meaning with differing ways to be made true. See also footnote 9.

²⁵This technical usage seems to stem from Langacker (1990), who discusses (p. 61) a person looking at the night sky, describing what s/he sees as a *constellation*, a *cluster of stars*, as *specks of light in the sky*, etc. He says "such expressions are semantically distinct; they reflect the speaker's alternate construals of the scene. . . .".

Cognitive Linguists as being different meanings for one and the same piece of language; the apparent reason for this being a belief that the semantic interpretation of some linguistic item is *directly* a particular mental structure without any intervening level of representation that might itself describe or be true of many differing values, and that these mental representations are *specific*. (Perhaps not totally specific, one presumes, but certainly more specific than Objectivists think their non-mental concepts are.) Objectivists naturally find this viewpoint mysterious and hard to take seriously, thinking that this is just a case of one and the same piece of language having different ways that it could be made true—that is, as being a case of generality. But here are a few quotations to give the flavour of the view, according at least to the types of Subjectivists we are canvassing.

Expressions do not mean; they are prompts for us to construct meanings by working with processes we already know. In no sense is the meaning of [an] . . . utterance ‘right there in the words.’ When we understand an utterance, we in no sense are understanding ‘just what the words say’; the words themselves say nothing independent of the richly detailed knowledge and powerful cognitive processes we bring to bear. (Turner 1991, p. 206)

Thus, there is no “general meaning” for a word, only the specific meaning as it is used on some particular occasion.

In [some] cases, varying content appears to originate in background situations. When representing *sofa*, for example, people might include information about sofas occurring in living rooms. Similarly, when representing *truth*, people might include information about truth being relevant in courts of law. In general, considerable amounts of information in conceptual representations appear to describe the background situations in which concepts occur, including settings, related objects, other agents’, actions, and variety of internal states (e.g., goals, evaluations, affect, motivation, reward, mentalizing). . . . The concept of *chair*, for example, is likely to be associated in memory with background situations for kitchens, living rooms, classrooms, theaters, jets, and ski lifts. . . . Sitting in a jet chair, for example, is associated with a different setting, actions, and internal states than sitting in a living room chair. . . . As a consequence, when a concept is processed in a particular situation, the situation activates related background information that enters into task performance. In other words, the extrinsic properties active for a concept also vary across situations, not just its intrinsic properties. (Barsalou 2017)

And so each particular situation uses the word in the sense appropriate exactly to such a situation, and not in its general sense.

Objectivists complain that the Subjectivists are so enthralled with the *particularity* of an experience that they have blinded themselves to the fact that language is full of generality in the description and reporting of that experience and refuse to acknowledge that this is precisely what happens on particular occasions. Indeed, they say, that’s the whole point of general terms. You use X to say something, then it is what X is true of that counts. . . not some hidden (or even public) idiosyncratic interpretation of X that you might have. If I say that I met someone on the train yesterday, an Objectivist would claim that it is just silly to think that I “really” meant that I met a woman, just because I know that it is true that the person is a woman! Objectivists think the Subjectivist theory is completely bewitched by the view that what is being communicated is just *exactly* what one is aware of and what one’s par-

ticular background knowledge picks out in some specific case. That's just plain false, they say: it depends *exactly* on *what was said*.

The **what did s/he say?** test shows that it can't depend on what the speaker happens to believe, even when what the speaker believes *is* true: suppose a speaker says *Connie saw a cat yesterday, but Delia didn't*. And let us suppose that the speaker knows it was a black cat that Connie saw, and that she saw it in regular daylight, with her normal eyes, while paying attention. So the speaker knows that Connie is aware that she was seeing a black cat.

Notice now that what I just said in those circumstances—that Connie saw a cat yesterday but Delia didn't see one—certainly wouldn't be true if Connie saw a black cat and Delia saw a tabby! But such a circumstance *would* make that claim be true if the sentence were to mean what Subjectivists claim: that Delia didn't see the same black cat that Connie saw (or same colour of cat that Connie saw), as seems to follow from the Subjectivist position under discussion. For, Delia can surely see a tabby while she is not seeing any black cat! This shows, using the **what did s/he say?** test that the speaker did *not* mean that Connie saw a black cat—despite the fact that the speaker knows it was a black cat and despite the possible fact that the speaker's mental image of the incident included a black cat, and that s/he would assent to the claim that Connie's seen cat was black, and so on. That simply is *not what was said* and hence not what was meant on that occasion.

The way to get the correct intended meaning is to make the *a cat* not be any more specific than "being a cat". Add any more and the sentence could be made true if Delia didn't see this extra material, even though she saw a cat. The view that every meaning is as specific as the speaker happened to have in mind would make lying or disseminating impossible. Diplomacy would be impossible. Romantic relations would be impossible. To Objectivists, it's as if Subjectivists about language have a morbid fear of things—even mental things, even linguistic things—that are not tied down to specifics.²⁶ One wonders what they might think about the concept SPECKLED HEN and the number of speckles it might have (Tye 2009).

On a more conciliatory note, one might wonder whether the following is the root cause of this dispute on generality. The truth-conditions of the *language* are general, even when the relevant concepts are particular in their nature. But which should be called the meaning? Objectivists point to the world, using arguments of the sort just canvassed. Subjectivists point to the mind. Given the radical differences in what is being "pointed to" by the two theories, they can't both be right about the same thing. So—is one right and the other wrong? or are they just talking about different things, calling them both 'meaning'? Or—as we explore in Sect. 7.3—is it that meaning has two different aspects, while still being one item?

²⁶I'm speaking generically, of course.

6.3.2 Grounding Concepts

We continue now with the general issue of subjectivity, and the Subjectivist's apparent commitment to meaning—and to thought more generally—as being “private” and within the mind of a single person. This is obviously a position that none of our Subjectivists²⁷ wants to hold in such a blatant form, for the reasons that I have given above. The standard response by our Subjectivists is to invoke some notion of “groundedness”.²⁸

There are many such notions in their literature,²⁹ sometimes going by other names, such as ‘situated’, ‘embodied’, and ‘enactive’—with subtle differences in application. (‘Grounded’ is used both as a general cover term and also more specifically to cover a notion of “gained by experience”; ‘situated’ is often used to talk about how a concept is activated in a particular situation; ‘embodied’ seems to be used to avert to the fact that we are physical beings and thus live in the physical world; ‘enactive’ is used for a more neuropsychological causal pathway. I’ll continue with ‘grounded’ as a general term for all these.³⁰) The idea behind them all is to give the otherwise private meanings (and other mental items) a way to “touch reality”. (Perhaps a way to avoid the “Brain in a Vat” scenario to the effect that since we never have any interaction with “reality”, we can’t know or talk about it (Putnam 1981).) Employing one of these notions is pretty popular in all the forms of Subjectivism about meaning we have been looking at, but while at least some varieties might seem to Objectivists to be plausible, the high degree of “fuzziness” in the articulation of how it in detail works—as well as the imaginative terminology that has been invented to describe it—has not met with much in the line of conversion of Objectivists. To many such un-convinced Objectivists, it seems no more informative to be told that such meaning concepts are grounded, situated, embodied, or enacted, than it is to be told that inter-subjective validity is guaranteed by “our shared form of life”.

The general reception by Objectivists of the notion of grounding in this form seems to be: Subjectivists can talk and talk about how one’s mental concepts and other meaning-related mental structures are ‘grounded’ in experience, ‘situated’ in context, ‘embodied’ in one’s perceptual/motor systems, or ‘enacted’. But so long as this process—whatever it is—allows for differing results on the mental items for different people or the same person at different times, there has been no explanation of how communication of mental structures to one another can succeed.

²⁷Recall that I am speaking generically!.

²⁸A general overview of the area can be gleaned from (Shapiro 2014) and (Pecher and Zwaan 2005), or at a higher level of generality in (Wilson and Foglia 2016).

²⁹A negative view of the many different notions called “groundedness” is offered in (Wilson 2002). A positive view that nonetheless details many different meanings of the notion is in (Rohrer 2007, pp. 28–31).

³⁰See (Barsalou 2010, p. 619): “‘Grounded cognition’ reflects the assumption that cognition is typically grounded in multiple ways, including simulations, situated action, and, on occasion, bodily states.”.

6.4 *Compositionality Cuts both Ways*

Although the privacy topic is, I believe, the root cause for the differences between Objectivists and Subjectivists, the literature from either side does not seem to dwell on this topic. Instead, we find other disputes that are tangentially related to this, but this root cause is not often mentioned. For example, Objectivists have long pointed to a set of “arguments” that are supposed to establish compositionality³¹:

Argument from understanding We can *understand* an infinite number of novel sentences, so long as they employ words we already understand. We understand sentences and other combinations that we have never encountered. So, language must be ‘compositional’: it must start with a finite stock of words/morphemes with their meanings, and put these together in a finite number of different ways, but using an unlimited recursive method to arrive at the infinite number of understood sentences.

Argument from productivity/creativity We can *create* new sentences that we have never heard or used before, and we know that they are appropriate to the situation in which we use them. This can only happen if language is ‘compositionally’ organized, so that we learn some finite base of words and rules, but know how to combine them recursively so as to produce totally new descriptions with the intended meanings.

Argument from learnability We are finite creatures who are exposed to a finite amount of information concerning our language. Nonetheless we can *learn* a system that is capable of infinite expression of meanings. The only way this can happen is if what we learn has a finite basis of terms and rules, but the rules themselves allow for arbitrarily complex ‘composition’.

As I see it, there are three threads interwoven in the arguments:

1. that language is something special (infinite, or novel, or creative, or whatever) in its ability to express meanings,
2. that people manage to use/learn/understand language despite their being “finite”,
3. that one (the only known?) way to do this is if language exhibits a semantically compositional framework.

It is important to keep these arguments in mind when reviewing the Objectivist position. Note in particular something we will return to in Sect. 7: Thread (2) of this general position cites certain abilities or outcomes of mental operations in language speakers. So whatever else is in the Objectivist position about meaning being “in the world”, at least in their only direct arguments for compositionality, there is a presumed subjective aspect.³² And although many Subjectivists try to establish that the

³¹Well, it is admitted by all that these are “arguments to the best explanation” and do not *logically* entail the conclusion of semantic compositionality.

³²In the linguistics literature there is also, to be sure, some argumentation concerning the proper form of a linguistic theory; and in that milieu, semantic compositionality is praised for its clarity and transparency. But these “methodological” (aesthetic?) considerations don’t really count as direct arguments in favor of the position, in my mind.

information present in the world impinging on language learners/speakers will allow for the Thread (2) to occur without compositionality (even though Thread (1) is correct), it seems to me that they would be on more secure ground in arguments against Objectivism were they to focus more attentively on the Objectivist's "missing link" (that these arguments can't succeed in securing (3) if the underlying presupposition of Objectivism is correct).

I have just argued that the Objectivists have painted themselves into a corner by employing as their only justification for (Functional) semantic compositionality an argument that presupposes the falsity of Objectivism. But what is the alternative, if one wanted to somehow include the Subjectivist presupposition? Two thoughts come to mind, and both can be found in the Subjectivist writings (mostly in the Cognitive Linguistics works). However, they are often stated in such a way that it is not clear which of the two thoughts is being promoted. And to my mind, this is because the two can be seen as shading into one another along a particular dimension, which I will try to elucidate. As I see it, one of the two understandings is an impossible position for anyone to hold, but the other one is not at all impossible. But given that the two ends of this dimension shade into one another, a lot of the argumentation has been against the impossible end and it has been allowed by some to impugn the other end by a sort of guilty association. Even granted that it is difficult to stop the "slippery slope" from the one end to the other, I will try to emphasize an underlying difference so as to take a step in rehabilitating the reputation of the group of possible theories that inhabit the one end.

One end of the dimension along which these theories vary is given by such theorists as Schank (e.g., 1972) and Anna Wierzbicka & Cliff Goddard (e.g., Goddard and Wierzbicka 1994; Wierzbicka 1996; Goddard 2002). In these theories there is a finite set of mental primitives in terms of which all other concepts can be defined. As I explained in Sect. 2, theories that can define all larger wholes in terms of their parts are Atomistic in nature, and so this version of Subjectivism is Atomistic. As I also explained in that section, theories that are Atomistic can often be compositional: at this end of the dimension under discussion here, we see that it is maybe possible to see the meaning of any defined concept to be a function of the meanings of the parts of that concept, and so on, down to the basic primitive universals. But one might relax the idea of a small set of primitives and allow more and more primitive concepts (or primitive word-concepts). In fact, one might allow *every* word to be such a primitive, thereby allowing *every* word-concept to be used in defining the others. Depending on how this is carried out, it may be either Wholistic or Atomistic, and the resulting semantic theory might therefore be either holistic or atomistic. In the former case it cannot be compositional, as I pointed out in Sect. 2, since holistic theories embody a type of Wholism, and by definition a wholistic theory is not compositional. But in the latter case, it might be... depending on the details of the theory.

One version of this more inclusive picture is offered by Langacker (1987), and I interpret it as Atomistic even though it makes all word-concepts available as possible members of the meaning of any other word-concept. (Langacker is being cited here with approval by Evans 2009, p. 61):

What Langacker appears to have in mind is that the semantic material – informally the meaning – associated with a lexical form, i.e., a word, relates directly to the contents of conceptual structure. In principle, this conceptual structure relates to a diverse and sophisticated body of non-linguistic knowledge The meaning of *uncle*, on this view, is potentially a function of the vast body of encyclopaedic knowledge we have of what it means to be someone's uncle. In addition to the specific relationship holding between the child of uncle's sibling, this also includes detailed knowledge relating to marital relations, familial relations, the social status of uncles, the types of behaviours associated with uncles, as well as individual knowledge any given individual may have with respect to uncles they have known. Yet while this knowledge is encyclopaedic, it is for Langacker part of semantic structure, i.e., directly encoded by a lexical form. Langacker's argument is that there is no principled way of separating putative linguistic from non-linguistic semantic representation.

This sort of theory of the meaning of word-concepts is in the “middle” of my dimension, and despite that, I would wish to view it as being Atomistic even though it allows for meanings to form “circular” definitional chains due to the large amount of material allowed in the meanings of a word. At the far end of the dimension, maximally different from the Shank or Wierzbicka/Goddard theories, and forming the Wholistic version of these types of theories, are so-called conceptual role theories.

Conceptual Role theories can be found in very many versions of Cognitive Linguistics. Like the Langacker version I just outlined, these theories also do not have a small number of primitive concepts, but rather all word-concepts are allowed to appear in the definition or explanation of any other concept. What sets conceptual role theories apart from the Atomistic Langackerian theories, and makes them be Wholistic, is their employment of *semantic liaisons*. A semantic liaison is an implicative connection that one concept might have to another concept—typical liaisons are ones of semantic inclusion (e.g., that being a tiger *implies* being a feline), contrary semantic entailment (e.g., that being a mammal is *incompatible* with being cold-blooded), compatibility relations (e.g., that one object can be both a brother to someone and a cousin of someone). More generally, in conceptual role theories *all* such implicative connections that a concept A has are inherited by any concept that includes concept A as part of its definition. Thus, if concept A has concept B as a part of its definition, then concept A also has all the implicative relations that B embodies. It is usually thought that this means that every concept has all other concepts and their implicative relations as part of its meaning. So, the meaning of a lexical item will be the concept it occasions *plus all the “liaisons” it has with other concepts*. Here a liaison is seen as either some meaning-entailment or (less strictly) some other association that the concept has with other concepts (that is, all the encyclopedic associations that a person might have with respect to the concept plus *their* liaisons). The meaning then is seen as the collection of all these liaisons. And of course, that is true for all concepts, including for the ones that form the liaisons with the first concept. In other words, meaning now becomes a feature of the entire set or body of concepts, and insofar as one can attribute meaning to any one concept, it is

most accurately said that this concept's meaning is its place in the entire network of liaisons among the concepts. Or, in the terminology mentioned above in Sect. 2, this is Semantic Holism—the competitor to Functional Semantic Compositionality. But it is not just a competitor to Semantic Compositionality, it is a competitor to any theory in the Atomistic camp... even the Langacker-style theories surveyed just above. In my mind it is important to keep these Atomistic vs. Wholistic versions of this style of theory clearly separated, when we wish to evaluate candidates for a Subjectivist theory of meaning.

Many theorists—and not just the Objectivists—find meaning holism wanting,³³ since according to this view, it would be very likely that different people will have had different language-learning experiences, and therefore according to semantic holism will have different liaisons among their lexical items. In turn, this means they do not speak with the same meanings associated with their words as one another. So, it follows that they do not mean the same thing as one another even when they utter the same word, phrase, or sentence. But then, when they (think they) agree with one another, or disagree with one another, they are in fact not doing so. Any evidence one has that you disagree with your conversational partner is, according to the theory, equally good evidence that you are talking about different things, and not disagreeing at all. And should you (appear to) *agree* with someone whom you don't like, that is better evidence that you are talking about different things, since you are independently motivated to disagree with him or her.

It is also often pointed out that, not only is communication with others an impossibility according to the theory's own features, but also one cannot communicate with oneself from one year to the next, one day to the next, one hour to the next, ... For, in the intervening time period the person would have new experiences and (except in very special cases where the person is comatose) these will impact the class of liaisons that impinge on the meanings of his lexicon. Although one *thinks* one remembers that the bark of eucalyptus trees peels off yearly, there is no justification for this, according to the theory. Over the year when you last had that thought, many of your liaisons have been altered—or at least, it is most likely to have happened—and thus your last-year-thoughts do not necessarily have *any* connection with this-year-thoughts.

Indeed, a Subjectivist-holist can't change his mind about *anything*! For, the very act of getting new information makes it be a different thing, and hence it is *not* a case of changing one's mind about concept X. The old concept X is no longer there to have a new opinion about. Not only is it impossible to disagree with another person, as I remarked above, but it is impossible to disagree with the past. We think that we have learned that Anaxagoras was wrong when he said that the unevenness of the moon's surface is due to the mixture of earthy matter with cold. But Subjectivist holism is committed to claiming that we are in fact not at all disagreeing with him. We can't disagree. We can't agree either. We've just changed the subject.

³³ (Fodor and Lepore 1991) is perhaps the best-known opponent. It seems that one author is a Subjectivist and the other is an Objectivist about meaning.

At this point it is tempting for our Subjectivist-holist to retreat a bit from the view. But it seems that the relevant retreat would be to a position akin to the Langacker-style Atomistic Subjectivism of very complex, but separate, concepts—which is *not* holistic. And so it might be a candidate for the Subjectivist portion of a two-tier theory imagined in Sect. 7.3.

7 (Some) Desiderata for a Two-Tiered Semantic Theory

A two-tiered semantic theory is one that alleges an accommodation of both the Objectivist and the Subjectivist theories, or at least of their main aspects or main contentions concerning meaning. Some versions of such theories might keep the two tiers separate, in the sense that neither one has any particular effect on the other, so that the features of the subjective (mental) never affects anything about the features of the objective, and correspondingly, the objective features never affect the subjective ones. We might, for example, follow a once-standard usage and call the subjective features that we have been discussing the *narrow meaning* or the *narrow content* of beliefs and meaning; and call the objective features the *wide meaning* or *wide content* of language. We could then form a theory of narrow content of mental concepts and its effects on, as it may be, beliefs, desires, and actions. And a different theory of wide content for linguistic items that gave an account of the truth and falsity of utterances.

But even if such a disjunctive two-part theory could be successful, each part in its own realm, it wouldn't be very satisfying to either the Objectivists or the Subjectivists, who each think that their theory accounts for some very important aspects of the unified activity or concept of meaning. What is desired is an "interactive" theory that shows how the two are related. And to do this requires some point of contact: a *nexus* of interaction. This nexus needn't be anything especially exotic; it need only be some item that both theories can agree upon, but which leads each theory in its own way. I describe one possible nexus in Sect. 7.3.

There is a worry here, though, when one tries to integrate these two theories. (Lewis 1970, p. 19) puts the worry thus:

I distinguish two topics: first, the description of possible languages or grammars as abstract semantic systems whereby symbols are associated with aspects of the world; and, second, the description of the psychological and sociological facts whereby a particular one of these abstract semantic systems is the one used by a person or population. Only confusion comes of mixing these two topics.

Partee (1979) has also discussed some worries. She describes a group of phenomena that she thinks cannot be handled by (what I call) Subjectivist theories. And in considering what a Subjective-theorist might say in reacting to this challenge, she considers three general attitudes they might take, but thinks that they all are incapable of handling these (propositional-attitude-related) phenomena in a satisfactory manner. In the end she despairs of finding any satisfactory solution. She calls the

three approaches (i) the idealized approach, (ii) a realistic-egocentric approach, and (iii) a realistic-nonegocentric approach; and it is useful to appreciate these options and the shortcomings that they give rise to. The option (i) seems to require a systematic understanding of what the “non-ideal” mechanisms are, and how an idealized theory can nonetheless account for them by means of some type of “contextual or performance factors”, in the same way that friction and the like are contextual factors affecting ideal laws of motion. She thinks this approach would not realistically be possible in the realm of meaning. Approach (ii), she thinks, requires figuring out one person’s (one’s own?) psychological performance factors, allowing one thereby to have a semantics for one’s own mental economy. But the remaining part of such a theory then requires attributing those factors to other speakers, and this seems like an implausible direction to go, given the difference in relevant experiences we all have. The third approach is, in her estimation, just “giving up” on the problem. It amounts to saying that, unlike option (ii) where we discover our own subjective theory and attribute it to others, we instead just say that *there are* some such factors and every person employs their own, whatever they may be. We merely “add variables over psychological mechanisms incorporating performance factors and add variables over psychological interpretations of individual lexical items”, and then existentially quantify over these variables. She thinks that this is just a retreat from trying to account for both Subjectivism and Objectivism in a single theory. It amounts to saying that since we don’t know what the relevant features and factors are, we just put existential quantifiers over them.³⁴

(Partee 1979) closes with this remark

What I have tried to suggest is that the [cognitive] linguist’s concern for psychological representation may be relevant to every semanticist’s concern for an account of propositional attitudes. So far I don’t see how to achieve either goal; my only positive suggestion is that a good theory might be expected to achieve both at once.

I do agree with Partee’s “positive suggestion”. But I also don’t wish to engender Lewis’s “confusion by mixing the two topics” of Objectivism and Subjectivism; and rather, I would want to show that they each *can* play a role in a joint semantic theory.

One aspect of Lewis’s admonition not to mix the two types of theories is the conviction that attempts to provide the two tiers both from *within* one side of the Objectivist/Subjectivist divide is not the right way to go—this is where the confusion comes from. I will try to demonstrate why I think this in Sects. 7.1 and 7.2. A part of the story will be that, coming from within just one side of the dispute, there is no common *nexus point* upon which the two can agree. It seems to me that the way to avoid Lewis’s worry is to show how there can in fact be a common nexus, and from this common point of agreement each of the theories goes its own way and does not necessarily affect the inner workings of the other. Below, in Sect. 7.3, I mention some considerations that may alleviate the worry that with only a nexus for connection, the two theories could differ in all sorts of unimaginable ways.

³⁴Partee also thinks of Lewis’s positive approach in (Lewis 1970) as being like this.

7.1 Objectivist Two-Tiered Theories

Some apparently two-tiered theories are actually theories *within* an Objectivist framework. Kaplan (1989), for instance, distinguishes ‘context of utterance’ and ‘circumstance of utterance’, as well as *character* and *content*. A way of putting his distinctions is that the *character* of an utterance is “the meaning without taking into account the context of utterance” or perhaps “the meaning in the language”. In any particular context of utterance, the values of the various pronouns, demonstratives, times, etc., that occur in the utterance are fixed, and then the evaluation of the character at that context will yield a *content*. Now using this content, we can evaluate it at a particular “circumstance” (maybe a particular possible world, such as the actual one) and the result will be a truth value. Kaplan’s motivation for the distinction between character and content involves the interpretation of pronouns and other indexicals (perhaps extended to predicates and the like), and not, generally speaking, the Subjectivist-oriented concerns about the structure of concepts and how they lead to beliefs, desires, and other features that guide behaviour.

So this picture plainly has both tiers working on the Objectivist side of meaning: once the values of the contextual variables are determined, the resulting semantic content is pretty much what one would expect from an Objectivist—a function on a possible world or “circumstance” to a truth condition. The semantic value of the character of a sentence is a function whose value at any context is the propositional content at that context. That is, a function that takes contexts as arguments and returns content, content being a truth conditional statement, namely a function from contexts and circumstances to some truth value. (See Westerståhl (2012) for how to make Kaplan’s theory be compositional.)

Another theory that comes from the Objectivist side is Frege’s, although perhaps the undeveloped Subjectivist side perhaps *could* be developed in some non-Objectivist way. Frege’s semantic theory of meaning (*Bedeutung*) and sense (*Sinn*) contains a subtheory which has certain explanatory powers relevant to a two-tiered account, even though it has its roots in an Objectivist framework.³⁵ In the objective, physical realm are the usual members of the physical universe of ordinary (and scientifically discoverable) objects. In the objective, abstract realm are not only such items as numbers, but also concepts and functions—taken to be the referents of predicates, relations, and operations on objects. Singular terms of language mean (that is, *bedeuten* or denote) the ordinary objects and the number-like items. Monadic predicates mean the concepts, relational expressions mean the relations, and function terms mean the functions—all members of the abstract realm.

A person can *grasp*—a technical term—these items in the abstract realm and thereby consider and possibly understand them (in the ordinary senses of these terms). But this is always done under some *mode of presentation* of the thing thus grasped. Frege insists on this way of putting how objects are grasped under a mode of presentation because of his realization that one needs somehow to distinguish cog-

³⁵Well, taking Frege’s view of the “abstract”, non-physical, non-mental, “third-realm” repository of concepts and thoughts as being a version of Objectivism.

nitive states of a person when the person has two names for the same object. The sentences in (3) obviously have different cognitive statuses, and yet the two names mean (denote, *bedeuten*) the same object:

- (3) a. Allen Stewart Konigsberg is Woody Allen
 b. Allen Stewart Konigsberg is Allen Stewart Konigsberg

The mode of presentation ALLEN STEWART KONIGSBERG designates a person and so does the mode WOODY ALLEN. But being told that the person designated via ALLEN STEWART KONIGSBERG is that person who is designated via ALLEN STEWART KONIGSBERG is cognitively trivial, whereas being told that it is the same person as the one designated via WOODY ALLEN is (or can be) informative.

Frege does not give any account of how the subjective world ought to process this “grasping”, nor what would happen in the subjective world that causes us to recognize when modes of presentation are identical, or how certain conceptual features would lead to assent or belief or action. Nonetheless, this picture has guided those Frege-friendly Objectivists who want some (at least minimal) psychological account of (at least) the informativity of identity statements. Frege’s type of formal move seems relevant to having a two-tiered semantic theory: the Objective theory would see no difference in the semantic values of (3-a) and (3-b), but the Subjective theory would make it important.

7.2 *Subjectivist Two-Tiered Theories*

Subjectivist theories that attempt to also account for the “objective reality” aspect of language all have the feature that they replace the Objectivist slogan of *language* representing reality with the slogan that (some aspect of) people’s *mental life*—for instance, their mental concepts—represents reality. It would seem to take some fancy footwork to justify the replacement of an unexplained notion of representation with another unexplained notion also to be called a representation.

One way that some Subjectivists have responded to the challenge that they need to have two distinct tiers was to take the viewpoint that they can use the distinction between people’s judgments of membership in a category and their judgments of the typicality of membership in that category. The membership judgment could perhaps fill the role of the objective-tier membership in that category. But from the Objectivist point of view using the judgements of membership as a substitute for “really being a member of the category” is the sort of idea that will give rise to Lewis’s “confusion”. At the best it can be *evidence* for membership in the category, but is clearly “contaminated” (the Objectivist claims) by all sorts of other biases. We can see that this can’t really be the Objectivist’s “fact about reality” by considering some of the actual experimentation that tests this. For example, the fact that subjects will judge a middle-height person as “both tall and not tall” in strong preference both to “tall” and to “not tall” (Alxatib and Pelletier 2011) does not in the least tend to support any theory about the reality of people manifesting contradictory properties.

Attempts to separate the two types of judgments (membership and typicality) and use one as a proxy for “reference” just miss the point of Objectivism. Although we can see this idea as endorsing a two-tier theory, we can also see that the use of “membership judgment” as if it were real membership is not something that Objectivists would agree to. And this is because judgments of whatever type just aren’t the sort of “neutral” nexus point that true two-tier theory requires.

Classic examples of this sort of “only apparently two-tier” theory from the Subjectivist side, provided by the group of theories that separate the notions of typicality from membership in a category, are (Smith et al. 1974; Osherson and Smith 1981; Smith and Medin 1981; Smith and Osherson 1984; Armstrong et al. 1983; Bourne 1982), among many others that are more recent (e.g., Winter 2017). In this arena, the main emphasis seems to be an argument as to whether there really are two different processes that people engage in. For example, Armstrong et al. (1983) demonstrated that typicality judgments are given to concepts whose membership was categorial by anyone’s standards, for example, odd versus even numbers. Such a result is then taken by this group as evidence for a two-tier theory.³⁶ But notice, as I said just above, the presumption that this can be what Objectivists mean by “membership in a category” is just wrong. An Objectivist will insist that it instead is a person’s *belief* or *judgment* that it is or isn’t a member. So, this is not the type of theory that I wish to call a true two-tier theory. A part of the objection is that such theories do not provide any *nexus* of contact between the two realms that is agreed-upon by both. Here we see the Subjectivists take the *concept* of category-membership as the nexus; but in the Objectivist’s eyes, this just raises the issue of how such a mental conception be attached to the external world. In a way, this is the reverse of the Subjectivist’s critique of the Frege “mode of presentation” nexus as a way to solve the issue of co-referring proper names. A Subjectivist would not (or: not necessarily) agree that this is where the two realms should meet.

Earlier in this subsection I remarked on the “fancy footwork” that might be required to make replacement of one unexplained notion of representation by another seem like an advance. Chomsky’s “Minimalist Program” has sometimes been thought to exemplify this sort of footwork. I will display this by a few quotes from McGilvray (1998), who is here explicating the notion of meaning in the Minimalist Program. We will not pause to evaluate its accuracy as an account of Minimalism, but instead will use it as a further example of a two-tier theory of meaning that has both aspects on the Subjectivist side. The first point in McGilvray’s exposition is to argue that the “external world” cannot be relevant to meaning, on the grounds that it is not “near” where the linguistic computation is taking place.

[B]ecause [linguistic] computations must also be local – they must proceed over ‘neighbouring’ elements, or be local – practical necessity restricts it to items in the head. Extending the domain of a theory of a cognitive competence outside the head makes it responsible for far too much; a theory would have to cover so much that it would lose any serious chance of being constructible. This poses a problem for an intentionalist view of a cognitive capacity, one that insists on determining internal states by their relations to things outside the head.

³⁶The negative side of the argument has seemed to many to be settled by Hampton (1988b). See also Hampton (2017).

Intentional properties arise when an inner state or process is treated as a representation of something in the world, or as having referential or alethic properties.³⁷ (McGilvray 1998, p. 231)

Further to the thought that the external world can't participate in meaning, McGilvray later says

Although [linguistic structure] does not deal with the kinds of semantic issues that [Objectivists] have in mind (reference and truth), it does deal with meanings. Generally, I suggest that the domain of syntax includes all locally determined, intrinsic features of linguistic mental events. If so, syntactic sciences include any disciplines that describe and explain the elements of this domain. As we have seen, the locality condition excludes reference and truth. (McGilvray 1998, p. 243)

And finally, McGilvray tries to show how this “perspective” allows for an internalist viewpoint on linguistic meaning which nonetheless has certain “ties” with the external world by means of when language is employed “in the world.”

Chomsky's view of [the semantic value of a linguistic item] as “a perspective” was an adaptation for a general audience of the technical syntactically defined concept of [such a value] as “an interface”. He points to a way in which [these values] are both rich and anthropocentric, and suggests a way to conceive of how they work cognitively in “interpreting” the world: [these values] focus attention on selected aspects of the world as it is taken to be by *other cognitive systems* [my emphasis], and provide intricate and highly specialized perspectives from which to view them, crucially involving human interests and concerns even in the simplest cases. I have in mind by richness the idea that [these values] provide very fine-grained media for all sorts of enterprises, not only describing and explaining the world (although not as a science that uses vocabulary apart from natural language would), but writing and reading literature, chatting with friends, and so on. (McGilvray 1998, p. 256)

We note that the connection to objective meaning no longer counts as semantics, in McGilvray's presentation of Chomskian Minimalism. It instead becomes part of the way the (rather mysterious) entities that are the “intrinsic features of linguistic mental events” are processed by the non-linguistic mental realm. Appropriately, this would relegate both the Objectivist's semantic values and the Cognitive Linguist/Psychologist's meaning concepts to a non-semantic realm of (say) pragmatics.

This viewpoint then seems to have denied *any* nexus for Objectivists and Subjectivists to agree upon for the interpretation of language and meaning. What starts as the apparent promise of a method of accommodating both Objectivists and Subjectivists in the realm of meaning has become a theory that neither side will agree to.

³⁷Note McGilvray's use of ‘intensionalist’ here: it means that the mental item *is about* something else—something in the objective realm, for instance. But it could instead be about some other type of *mental* entity. This sort of intensionalism would be friendly to those Objectivists who want their representations to “be about items of reality”. Thus McGilvray is here arguing against Objectivism, and so is presumably attributing a form of Subjectivism to Chomsky's Minimalism.

7.3 Two-Tiered Theories with a Nexus

It is perhaps most clear from Frege's example of "grasping" that we would want the Objective Tier of our two-tiered theory to give an account of truth and possible truth (etc.) by means of features of the world (or of features of possible worlds that are objectively available in the actual world). And we would want the Subjective Tier of the theory to give an account of the way that the world has "presented itself" to an individual to account for aspects of that person's belief states, desire states, and intentional actions. (Some sort of a *Belief plus Desire entails Action* theory would give an example of the sort of Two-Tier theory I have in mind.) This is the sort of theory that Jerry Fodor seems to have had in mind during one of his earlier states of theorizing (Fodor 1980), where *wide* versus *narrow* content was introduced. But I would prefer a different method of "tying together" the Objective and Subjective than Fodor seems to have employed during this period of his thought on the topic.

From such a viewpoint, it also seems clear that the desired *nexus*—the place where the two conceptions of meaning meet—is the word, or more accurately, the *individual sense of a word*, for the sort of reasons outlined in footnote 9 of Sect. 3. On the Objectivist side of this theory, the meanings assigned to these senses (some sort of reference or function on possible worlds) are the atomic units of meaning that form the basis cases for compositional-like (atomistic) accounts of meaning. On the Subjectivist side of this theory, it is the individual understanding of the mental meanings of these senses that combine to form complexes which in turn can interact to form beliefs, give rise to desires, and produce actions. My argumentation in Sect. 6.4 was aimed at the conclusion that this subjective side also needed to be atomistic in its treatment of meaning—although perhaps not compositional.

One recent foray into this realm is taken in (McNally and Boleda 2017), although perhaps motivated by different concerns than the one identified here. McNally and Boleda are concerned with the fact that "in the absence of any specific context" some sentences show a specific and particularly pronounced understanding, whereas when enclosed in a wider linguistic context they can vary so much as to make the original "default" interpretation be difficult to see. They see two sources of information as available for a language understander: the referential and the conceptual. And the two sources may operate differently, the latter being especially sensitive to "affordances" provided by surrounding context. As they put it:

In the case of language, we take its connections to concepts and to the world to be distinct features, each of which facilitates a distinct process of concept composition. When the concepts contributed by two expressions in a grammatical phrase naturally suggest the ways in which they should be composed, and interlocutors avail themselves of such a suggestion, we can say the concept composition is *conceptually afforded*. Alternative, if specific, independently available information about the referent described by the phrase is used to guide the way in which the concepts in question are composed, we can say that the concept composition is *referentially afforded*.

This general picture is rather attractive, and especially in that it locates the relevant nexus to be linguistic items. (Although I would have preferred them to have

identified specifically that they were senses of lexical items.)³⁸ Language is—as it were—“neutral” in how it is to be understood, but different cues give rise to a distinction in ways of understanding, between a referential interpretation and a conceptual interpretation. Like the Frege theory, this seems to give us a way to picture how a two-tier theory could be developed, this time from a Subjectivist point of view. Below I will likewise gesture at a way of looking at a two-tier theory. I intend mine to more strongly separate the (Objective) referential component from the (Subjective) conceptual-connection component. But like McNally and Boleda (2017), I will just gesture at the theory and not give a lot of detail. In Sect. 8, I describe a number of phenomena that need to be dealt with in the subjective tier before any such theory can work.

In brief outline: I see the two-tier theory working somewhat as follows. A sense of an individual content word³⁹ means^o (*mean* in the objective tier) the sort of items favoured by the formal semanticists—individuals for singular terms, sets of individuals for simple nouns, properties⁴⁰ for adjectives, activities and processes for verbs, and so on as required by a truth theory of meaning (or as modified for a theory of truth at a possible world). They mean^s (subjective meaning) some mental concept (of the appropriate sort).

The objective tier is semantically compositional—or at least, is atomistic with some sort of assignment of values for meaning^o in a way that honours the general argumentation for atomicity of meaning. It would be immensely pleasing if the subjective tier could be similarly organized—that all the observed differences present in meanings of complex linguistic expressions could be traced to the differences between the meaning^o and meaning^s of individual lexical items. But it is a long-established result that lexical items do not form the only place where the difference in the meaning^s of complex concepts is to be found. They are also found in the ways that meanings^s combine.

I view the referential portion to be in fairly good shape in these regards (it being the focus of formal semanticists for many decades and all aimed at a similar “truth in the objective world” goal), but that the concept-portion needs to answer many topics that have not been dealt with in the Subjective literature (and hence can’t be directly incorporated into a two-tiered theory, yet). See Sect. 8 for discussion of what other further clarification of these matters is still called for.

The crucial point of a genuine two-tier theory is that there is some point of contact—what I have called a nexus—where each side of the theory can agree upon the features of that contact, but such that each side can decide on what aspect of these features their particular theory should feature. As I see the nexus, it resides in a sense

³⁸And one might also object to this version of a two-tier theory on the grounds that McNally and Boleda’s view of reference is questionable (they appeal to a distributional analysis of the closeness of a word to other words in a corpus, using a multi-dimensional vector analysis). But we will not dwell on that here.

³⁹I do not here wish to discuss the rather different workings of such non-content words as the logical connectives, the copula-like words, the quantifiers, the multi-faceted prepositions, and so on. Various comments about them are below, in Sect. 8.3.

⁴⁰Functions from individuals to truth values (Or: functions from possible worlds to functions from individuals to truth values.).

of a lexical item. Such an item will have two parts: a meaning^o and a meaning^s, corresponding to some objective features of reality and some relevant subjective properties of mentation. Not only are these properties of a lexical sense different, but what can be done with each of them differs, due to their differing properties.

One crucial feature to be investigated concerning these differences is how far the two theories of meaning can diverge from one another. Given that the meaning^o is dictated by the fixed features of reality, the question then seems to become whether the mental realm can veer willy-nilly away from the objective realm. As thus far described, the answer is yes—no restrictions have been placed on the mental representations. It is only by experimentation that we can find out the fine detail of the subjective realm. However, on an *a priori* level we might say that, since the meaning^s is supposed to give the agent's reason for action by the way it can construct the agent's beliefs, it seems that these beliefs (in the main, anyway) can't diverge too, too far from the Objective realm. If the beliefs are all screwy, the agent won't be able to get along in the real world—for long, anyway. And it will be likely that many actions and expressed beliefs will be at odds with others of theirs and of other people's. That is, when meaning^s conflicts in action-oriented ways with meaning^o, bad things can and will happen. And so a basic thought about meaning^s is that in an important sense they can't diverge much from meaning^o. Serious divergence between (say) the objective reality versus the subjective beliefs of the eating habits of grizzly bears might limit the gene pool to be just those who do not have a divergence here.

In less serious cases, for instance where a person's beliefs about snakes have SLIMY as a component, different outcomes might occur. If nothing intercedes, they just go on believing that snakes are slimy and their actions (such as snake-avoidance) is explained on the grounds that they have an aversion to sliminess. When it is explained to them that snakes *in reality* are not slimy, they may alter their SNAKE concept and change their actions; but they may also *not* accept such a claim and continue their previous actions; or finally, they may *claim* they accept the truth^o, and therefore answer that snakes are not slimy when asked, but *not* change their SNAKE concept and not change their other actions that involve snakes. All three of these reactions are possible. But the second and third reactions can't happen too often, since it would be evolutionarily a disaster were people not to change their internal concepts in concert with what occurs in reality. It is this force that keeps the meaning^s connected to meaning^o.⁴¹

Asher (2011) develops the general form of theory that I wish to promote. Central to his notion is that lexical items contain two parts: a contribution to objective semantics and what he calls a TYPE. The objective portion is especially simple, but the effect of TYPE is intricate and detailed, with many different aspects, including pre-supposition enablement, context sensitivity effects, sense transfer, metonymy, predicate transfers, discourse effects, explanation of *qua* phrases, dedicative phrases, and other very subtle meaning-topics.

⁴¹James Hampton (p.c.) remarks that evolutionary pressure may not always work this way. Religious beliefs may serve to strengthen social groups to greater cohesion (and success) while being relatively immune to falsification by meaning^o.

Here are some of Asher's introductory remarks on his vision:

The guiding idea... is that almost all words will have single and simple lexical entries. Words like the nouns *cat*, *lunch*, *book* or the verbs *kill*, *read*, and *master* denote simple properties or relations, and so accordingly the logical forms that specify their denotations are very simple. For instance, *cat* has the lexical entry $\lambda x \text{Cat}(x)$ ⁴²... If words have simple entries and make simple contributions to truth conditional content, they come with a rich amount of information about the TYPE assigned... within them. These TYPES will guide predication and be responsible for fine-grained differences in lexical meaning. When words are combined together to form clauses, sentences, and discourses, the TYPE associated with various terms will interact in complex ways. I will introduce operations of TYPE adjustment in response to TYPE mismatches between predicate and argument that correspond to the accommodation of a TYPE presupposition, or more generally speaking, the justification of such a presupposition.... The effects of these TYPE adjustments... is that the logical form for a clause will contain elements that are not present in the lambda terms for the constituent words themselves. Predication involves not only applying a function to an argument but also operations of adjustment corresponding to type presupposition justification. (p. 22)

By his notion of "TYPE adjustment" and "coercion" he includes a very wide range of examples that are familiar from the Cognitive Linguistics literature, such as

- (4) a. good lunch versus good children
 b. Mouse isn't very tasty unless you're a cat
 c. John started a cigarette/started a car/began the sonata/started the novel
 d. John liked the dress with the flowers/liked the garden with the flowers

And he mentions certain entailments as also falling under the sort of "coercion" he is concerned to capture, such as

- (5) John started the car *defeasibly implies* John started the engine of the car

In Asher's account, another aspect of the TYPE portion of lexical entries concerns how they interact with the TYPE-entries of other lexical items in a longer stretch of discourse. This interaction is governed by the underlying logic of the mental side: his adaption of classical type-theoretic logic to become a proof-theoretic semantics for the TYPE side of his theory, as I suggested in Sect. 6.2.

I've argued that TYPES are concepts, mind-dependent entities with fine-grained content. TYPES have an internal semantics that is given at least in part in terms of the rules by which they combine with other TYPES; they are proof-theoretic objects. TYPES can be associated with other TYPES – what we might think of as *traits*; each trait we associate with a TYPE is a constraint on the introduction rule for the TYPE.... TYPES as proof objects provide an internal semantics for natural language sentences and discourses that complements the external semantics given by ordinary intensions. We can even distinguish between different conceptualizations of the same physical object or of the same property, so in some respects the structure of types has the capacity to make finer distinctions in meaning than intensional semantics can. But the purpose of the internal semantics is not the same as that of the external semantics. Intensions are the soul of a theory of meaning – they are needed to determine truth, reference, and other *external*, semantic properties that link language to the world we talk about. TYPES and their adjustments are the heart of a theory of predication and

⁴²This lambda term is to be interpreted as "the feature of being an x such that x is a cat", or more simply, "the property of being a cat."

responsible for other properties of meaning. . . . This is semantics but it is an *internal* matter, something that speakers do “in their heads.” (pp. 44–45)

The one aspect I find missing from this otherwise compelling story is that there is no room here for individual variation of meanings. In particular, the mental realm—as captured by the TYPE dimension of the dual-aspect lexical items and carried along via the proof-theoretic semantics to full sentence-thoughts is the same for all people, and it furthermore (appears to) contain exactly the *objectively true* features of the encyclopedic facts in the word-TYPES when these are manipulated so as to give the meaning^s for the sentence-thought. It thus cannot deal with issues of false beliefs about items in the world; it will not give any account of why different subjects give individually different results on the many tasks about meaning that have been administered over the decades in cognitive psychology.

Although Asher’s theory is not described this way, perhaps the alteration I would advise is to make his theory be a theory for an *individual agent*, allowing this agent to have his/her own set of encyclopedic “facts” for every lexical sense.⁴³ Presumably there would also be individual differences in the underlying inference mechanisms of the proof-theoretic semantics also, another topic not envisaged in Asher’s picture.

But this now would leave us without any way to accommodate “real communication”, where the two conversationalists are “really” talking about the same (mental) things. While there can always be some check against “the objective facts” due to the objective portion of meaning in Asher’s theory, this will not help conversation that relies on having some mutual understanding of the TYPE-portion of meaning—the meaning^s.

I leave this problem hanging, as a very important topic still to be solved by any two-tiered theory. And I turn instead to other aspects that need to be dealt with in the subjective, TYPE portion of the theory.

8 (Some) Aspects of Current Theories of Mental Concepts that Require Attention

Objectivist theories of meaning are just fine—so far as they go, which is not very far when it comes to mental life. (And one might reasonably think this is a very major shortfall!) On the other hand, Subjectivist theories need some further work, even within just the mental side of meaning with which they’ve been concerned. This section is about a group of topics that need to be addressed more fully than they have been, in order that one can take the Subjectivist theory of meaning seriously enough to be included in the meaning^s portion of a two-tier theory of meaning. I mention three areas that seem to be “obvious problems” that need to be dealt with before it is possible to develop any deep mental theories that invoke concepts., including the

⁴³I might mention that Asher’s picture does not take individual senses as the basic items, but rather uses TYPE-coercion to generate the distinct senses as required in any particular sentence.

desired concept-oriented portion of any two-tiered theory of meaning of the sorts mentioned in Sect. 7.

The first area concerns the type of concepts (and the details concerning them) that such theories envisage, and it has several subparts. The second concerns a venerable puzzle that should be considered by anyone when they embark on proposing any account of meaning, whether Subjectivist or Objectivist. The third area requests Subjectivists to construct even a basic version of what is already present in Objectivist accounts. . . and then asks whether there is any reason to go to the trouble.

8.1 (Some) Messy Details for Conceptual Theories

I mention in this subsection a series of interrelated topics that seem to me to draw attention to the inadequate attention that is paid to the vast majority of linguistic data by theories of concepts and conceptual combination. These phenomena are seen as central in Objectivist accounts, but which seem not to figure in most Subjectivist accounts. This holds, I think, not only for the psychological theories that postulate concepts to be tested against certain behavioural responses by language users (where these theories don't pretend to test a full range of linguistic phenomena), but also for the cognitive linguists who are very aware of many of the deep and peculiar intricacies of natural language but who nevertheless use the overly-simple results from the psychological theories to account for them. (Sometimes just assuming that "more of the same" will accommodate the intricacies.)

I write this from the point of view of an Objectivist, formal semanticist. All the topics discussed below are thought to be adequately dealt with on the Objectivist side of a two-tier theory (at least by Objectivists⁴⁴). Presumably, Subjectivists would have their own lists of topics that Objectivists need to deal with before any two-tier account could be made to work. But someone else will have to provide that list.

8.1.1 Basic Problems with Singular Terms

Proper names and other singular terms do not seem to feature heavily in the Subjectivist theories, although it seems pretty clear that these theories would somehow wish to generate concepts in an individual's mind corresponding to them. But proper names and other singular terms have a number of well-studied peculiarities that seem to tell against mental versions of concepts for proper nouns.

As we have seen, Subjectivist accounts claim that concepts embody "encyclopedic knowledge" of the items of which they are concepts. Thus, even though *NATIVE OF AFRICA* is not part of the meaning of *lion*, it is a part of the concept *LION*, in the Subjectivist's view. The philosophers of language have advanced a number of arguments that have seemed to Objectivists to show conclusively that this cannot

⁴⁴Speaking generically.

be true for proper names. Although TEACHER OF ARISTOTLE might be part of the stereotype or encyclopedic knowledge one has of Plato, we could discover that it was all a mistake in the transcription of one ancient document, and that in fact Aristotle was never in Plato's Academy. Nonetheless, even after discovering this, when we use the name *Plato*—as I just did in the last sentence—I was referring to the same person as I always referred to. And, it is usually added, this could hold of *any* of the encyclopedic properties of Plato—indeed, of *all* the encyclopedic properties (Kripke 1980).⁴⁵

A theory of concepts should in one way or another be able to accommodate this finding. How can a singular term concept such as PLATO evolve so as to lose some, and possibly all, of the properties given by our encyclopedic knowledge?

Similar remarks could be made about other singular terms. Unless one wants to make the concept answering to *The current President of the United States* be the same as that corresponding to *The person I believe to be the current President of the United States*, one cannot include in the concept any stereotypical properties that are not essential. But if you *do* decide that these were the same concept, you could never discover that you were wrong! (Because they are the same concept, they said; so any change just continues to make them (it?) the same concept.)

8.1.2 Basic Problems with Singulars versus Plurals, and Quantified Noun Phrases

It is surprising that there is not more mention of the difference between concepts that correspond to singular noun phrases in natural language and concepts of the corresponding plurals. A survey of any of the literature will show that simple nouns like *dog* get associated with a concept that is multiply-instantiated—that is, this concept applies to (or represents) any dog. Intuitively it corresponds to a plural (or perhaps to an unmarked-for-number) term. But in either case it remains somewhat of a mystery how this can be employed to indicate a *particular* dog or to assert something true of a single dog, and at the same time to assert something that is true of some but not all the dogs. Clearly there needs to be some way to incorporate numerals and quantifiers in the theory. Surely we want to conceptually differentiate the sentences *A dog is in the yard* from *Two dogs are in the yard* from *Dogs are in the yard*. Yet it seems quite difficult to have this sort of modification of simple nouns when using only the type of conceptual combination that is usually described in Subjectivist writings.

⁴⁵This claim—that an object identified by a proper name *could* lack all the properties traditionally associated with the object and still be referred to by the proper name—follows from any version of the “historical-causal theory” of reference (of proper names). One intuition supporting the claim is that any of the initial identifying features in fact *needn't* have been true—they need only *seem* to be true to the initiator of the use of the proper name. And then those who adopt this name into their idiolect just assume these to be factual claims about the initial referent. But as this consideration shows, the truth needn't be a part of this assumed set of features.

Furthermore, natural language contains a myriad of different types of quantified noun phrases such as *all people* just as much as it does adjectival modifications of nouns such as *tall people*. This suggests that since conceptual combination is used to generate a concept TALL PEOPLE from the concepts TALL and PERSON,⁴⁶ something akin to it should be used to generate the concept ALL PEOPLE. But this would require a concept ALL, much along the lines of the concept TALL. And this has seemed to most philosophers an unlikely suggestion: quantifiers are higher-order, that is, they are properties of properties. Here ALL would need to be *a property of* PERSON, not *one of* PERSON's properties.

Without a theory that accommodates the quantifiers (and there are many more than just *some* and *all*), the theory of concepts is essentially unable to accommodate the vast majority of our language and our mental life.

8.1.3 Basic Problems with Noun-Oriented Concepts

It is striking, although perhaps explainable, that theories of concepts start by presuming concepts corresponding to certain lexical nouns: the concepts PET, CAT, MOUNTAIN, CABIN, . . . , and so on. As mentioned above, and as is ubiquitous in the literature, the history is that very much of this research program's time is now spent on trying to determine just what the nature of these concepts should be: feature lists, prototypes, stereotypes, schemata, dynamic whatevers, and so on. A test for correctness of the resulting theory is then developed, and it consists in how well "conceptual combination" works according to these tests. But the examples thereby considered for these tests, at least in the first instances—PET CAT, MOUNTAIN CABIN, CHERRY BLOSSOM, . . . —seem to be carefully-chosen noun-noun compounds, since they don't include things like PET ROCK, CABIN CAT, CABIN FEVER, . . . , for which there are intuitively plausible interpretations but where the method for arriving at these interpretations would seem to be in conflict with the methods established for the initial examples. The difference of interpretation between IRON GATE and WATER METER, obviously requires some distinction between "material substance of" and "recipient of" to formally yield for us a gate made of iron versus a meter that controls or measures water.⁴⁷

Furthermore, as presented in the literature, the various mechanisms for conceptual combination in the literature are—at least in the main and for the most part, as emphasized in (Hampton 2011b)—"symmetric": for example, one considers the feature lists from two prototypes and constructs some sort of new feature list corre-

⁴⁶Note the insertion here of some sort of plurality operator to form the combined concept! As I said, there needs to be some way for concept theory to generate the singular-plural distinction.

⁴⁷There are, of course, the hints from those Subjectivists who propose types of "metaphorical interpretations" and rely on magic for the formalism. And more plausibly, there are suggestions about how there can be underlying principles governing differing types of this sort of compounding. For example, there are nice principles stated in (Wisniewski 1997; Gagné 2002; Gagné and Spalding 2006; Spalding and Gagné 2008; Gagné et al. 2016), but I would prefer to see some sort of algorithm instead.

sponding to the combined phrase. From PET and FISH, one constructs a prototype for PET FISH. But as a symmetric procedure, this very same result should be generated from the phrase *fish pet* also, even though it is very difficult to give an interpretation for that noun-noun compound, as it occurs in natural language.

But matters are more complicated here, since some noun-noun compounds do in fact have interpretations for either order of the lexical nouns, but the interpretations are different: *grease gun* versus *gun grease*, *talk show* versus *show talk*, *truck delivery* versus *delivery truck*, *school teacher* versus *teacher school*, *flea circus* versus *circus flea*, for example. All this is one topic that should be addressed at the very beginning of any theory of concepts. Hampton (2011b) outlines various types of “non-intersective” conceptual combination, referring to some seminal work in (Wisniewski 1997). Word order (in English) gives rise to much of the “direction of interpretation”: the head of a noun-noun phrase tends to be on the right. Although this is correct (generally) for English orthography and verbal production, it is not at all clear why it should dictate how *the concepts themselves* should combine. Is it that the concepts GREASE and GUN can combine in both ways freely, but that one is associated with the linguistic phrase *grease gun* while the other is associated with the phrase *gun grease*? But now what is needed is some principle that determines why one of the concepts is associated with the one linguistic phrase while the other concept is associated with the other phrase. Is there anything at all intrinsic to the concepts that dictates one over the other? (For some thoughts on this dark matter, see Hampton 2011a.)

Of course, disjunctions of nouns occur in language, as well as conjunctions. So we would like a satisfying account of the type of concept one gets for *either a dog or a cat* and (more difficult, perhaps) *either an aggressive intruder or a figment of the imagination*.

Most theories of concepts also explicitly introduce concepts corresponding to adjectives, so that we have concepts like HAPPY, TALL, RED, etc. Now we can have conceptual combination to yield concepts corresponding to *happy pet*, *green cabin*, *red rock*, ...; but in these cases we *do not* have symmetry of conceptual combination—*pet happy*, *cabin green*, *rock red*, ... are not normally grammatical when the second item is an adjective, unlike (most) cases where they both are nouns. (In those cases the phrases are grammatical but may be semantically incoherent. This seems a natural thing to say, but one can wonder whether the Subjectivist account of concepts will allow a distinction between ungrammaticality and semantic incoherence.) This argues that concepts corresponding to adjectives have to be differentiated from those corresponding to nouns. And as a consequence, conceptual combination of concepts of adjectives would probably be different from that for nouns: the adjective phrase *rich*, *happy*, and *well-adjusted* maybe corresponds to a conceptually very different sort of concept from the concept corresponding to the noun phrase *men*, *women*, and *children*.

I do not suggest that *no* work has been done on these topics—disjunctions and negations were discussed long ago in (Hampton 1988a, 1997). But as I will mention in Sect. 8.3, a theory that explains the workings within the realm of concepts that is of a scope as broad as the Objectivist account still remains to be articulated.

8.1.4 Basic Problems with Verb-Oriented Concepts

One presumes that lexical verbs also correspond to mental concepts in this overall picture, and that since verb phrases can conjoin (*eats, drinks, and falls asleep* as well as *marries, works, breeds, and dies*); so presumably the corresponding verb-oriented concepts can conjoin. But as I will try to demonstrate in this section, it seems reasonable to suppose it will be a different type of concept combination than in the nominal side of things, because of the particular differences there are between the noun-concepts and the verb-concepts, as well as the higher degree of complexity that verb-concepts display.

Conceptual combination seems to have been designed to deal with what might be called “conjunctive (or: intersective) concepts”, even though the resulting feature lists (or whatever used to keep track of properties) is not a simple conjunction of the feature lists of the subparts. A part of the rationale for concentrating on this could be that English allows easily for conjoined noun phrases (*Bob and Sally, Men and women, Both small dogs and short snakes, All baseball players but only some Olympic high-jumpers, etc.*) And also that the interpretation of adjective-noun combinations has historically been seen as conjunctive (although modern formal semantics treats them otherwise). But it is not so clear to me how any similar technique will work with what might be called “disjunctive concepts”. After all, English pretty easily allows for disjunctive noun phrases: *Juan or Alice, Men or women, Either Americans or Canadians, etc.*⁴⁸

As remarked just above, verb phrases easily conjoin just as much as noun phrases do, but I claimed that the conceptual items corresponding to verbal items will require some different sort of combining method from the ones in the noun realm. Verb phrases can disjoin, like noun phrases can, for linguistic phrases such as *is either swimming or eating*. Another place where noun concepts and verb concepts appear to be different is in the case of linguistic conjunction. But given that the conceptual combining method for verb-concepts is different from that of noun-concepts, due to the inherently more complex nature of verbs, it seems likely that the method for forming disjunctive verb-concepts would have to be different from that of forming disjunctive noun-concepts.

A special difficulty for theories of concepts attends to those concepts that correspond to non-monic verbs. That is, to relational verbs—ones that take direct and indirect objects.) For example, the verb *kick* requires a direct object, and so any corresponding concept will require something corresponding to this. As with any verb-concept, there will need to be some notion of a subject of the kicking-concept. But in the case of transitive verbs, the corresponding concept will also require the notion of a recipient of the kicking. It seems that some as-yet unspecified notion of conceptual combination is required for this. (And a further unspecified one for

⁴⁸In the olden days there were various works on disjunctive concepts, both learning and using them (Furth 1963; Conant and Trabasso 1964; Seggie 1969; Bar-Hillel and Eifermann 1970; Trabasso et al. 1971; Newstead and Griggs 1983), but these have seemingly diminished in recent years, and the few more recent writings have focussed on reasoning with disjunctions. However, there are works by, especially, James Hampton on aspects of these topics: (1988a, 1997, 2007, 2012).

indirect objects.) As I have mentioned above, noun phrases are not only these simple lexical nouns nor do they merely include only the conjunctive and disjunctive nouns, but they also include proper names and quantified noun phrases. It is most unclear how the imagined-but-as-yet-unspecified-conceptual-combination that creates verbal concepts with direct and indirect object concepts really works, even in simple cases like *loves dogs*, but it seems quite a bit more difficult to imagine any such method of conceptual combination that would work with *loves all dogs* or with *loves some dogs* or with *loves many dogs*, or with *loves all dogs but only a few cats*, etc., and give the intuitively correct distinctions among them all. (And with indirect objects we can have verb phrases like *gave many children numerous gifts*. It is again most unclear how such complex concepts can be built up, especially since the underlying concept for the verbs and noun phrases are unspecified as yet.) A further related difficulty arises with those verbs that can also be used in an ergative manner as in (6-b):

- (6) a. Mary boiled the water
 b. The water boiled
- (7) a. Kim tripped Sandy
 b. Sandy tripped

Yet another difficulty that seems not often addressed is the difference that English (and other languages) shows between “stative” and “eventive” verb phrases. (*Knows Japanese* is a stative verb phrase whereas *is speaking Japanese* is eventive, reporting an activity. Other types of eventive verbs indicate achievements or accomplishments. One linguistic test is that eventive verbs allow the progressive (in English), whereas stative verbs don’t: **John is knowing Japanese* is not correct, but *John is speaking Japanese* is.) This distinction also brings up the difference between *ability* versus *activity* senses. And that is another difficulty for concept theories. When FLY is conceptually combined with TWEETY,⁴⁹ is the result a representation of the natural language *Tweety can fly* or *Tweety is flying*? There is a very large literature in the intersection of philosophy of language and formal semantics that deals with issues involved with this distinction, and these ought to be reflected in theories of concepts (e.g., Parsons 1990; Schein 1993; Casati and Varzi 2015).

8.1.5 Basic Problems About Other Noun- and Verb-Oriented Concepts

The fact that the various major categories of syntax allow for conjunctive and disjunctive combinations, and that the Subjectivist theory thereby requires some type of concepts corresponding to them, leads one to wonder whether there can be *negative* noun-concepts, or *conditional* noun-concepts, according to theories of mental constructions. Similarly, one wonders about negative or conditional verb-concepts.

Although negation in English can apply pretty easily to verbs/verb phrases, and to adjectival/adverbial phrases with only a slight increase in difficulty, they apply with

⁴⁹ Assuming that the concept theorists have a way of representing proper names as concepts.

rather more difficulty to noun phrases. Can there be a concept NON-CAT? Maybe it's more plausible to have a concept NON-STARTER? But then there are more complex noun phrases, and if concept-theories are extended to handle arbitrary noun conjunction/disjunction, it ought to be extended with negation so as to handle the concept associated with *short adults but not tall children* or *neither oil nor coal*, and the like. One wonders, then, just how negative noun-concepts are going to be handled. My own intuition is that it will require very different applications to the different sorts of concepts (that is, it will apply differently to different grammatical categories). It would be good to see a careful account that will accommodate all this.

Since English easily allows for conjunctive and disjunctive noun phrases, adjectival phrases, adverbial phrases, and verb phrases, and allows easily for negative verb phrases, adjectival phrases, and adverbial phrases, it seems that the usual “logical connectives” can apply to any of the usual linguistic categories other than nouns. But then one wonders about non-standard connectives like conditionals: English allows conditionals in all the linguistic categories (although it is rather more stilted in the noun case): *is happy if he is eating* (conditional verb phrase), *if not tall then at least not short* (conditional negations and adjectives); *morally if not legally* (conditional adverbial phrase), *crocodiles, if not all reptiles* (conditional noun phrase, with negation). So it seems that there needs to be some sort of conditional-forming conceptual combination. Needless to say, this doesn't at all look like the usual conceptual combination.

8.2 *Concepts, Sentences, and Bradley's Regress*

We observe that there has not been any account of the interpretation of *sentences* given by Subjectivists. That is, there is as yet no account of a (mental) *judgment*. Even granting a singular concept SANDY and a verbal concept SWIMS, there is not yet any judgment that Sandy is swimming.⁵⁰

This is related to another question that concept-theorists seem not to have addressed about an old philosophical puzzle and to a more modern apparent paradox. The puzzle is: how can mere *words* combine to form a *sentence*? That is, what is the difference between a list of words and a complete sentence? How can it be that a list of words becomes something different—something that expresses an entire proposition? All sentences, even very simple ones, give rise to this puzzle: Consider

My list of words:

[Word #1] Kim

[Word #2] runs

My Assertion:

Kim runs.

⁵⁰Or that Sandy *can* swim, depending on the account given as an interpretation of the concept SWIM.

As various writers in the past have claimed, the former is “inert” and “is a mere congeries”, whereas the latter actually “does something” in making an assertion that describes a feature of Kim.

Note that a structurally identical puzzle arises if ‘Word’ is replaced by ‘Concept’ and ‘My assertion’ is replaced by ‘My judgment’. Both Objectivists and Subjectivists are subject to this puzzle, and they both should give some sort of answer. However, Subjectivists seem not to have thought of a difference between the noun-phrase conceptual combination {SANDY; SWIM} and the sentential conceptual combination {SANDY IS SWIMMING}. Nor has it seemed to occur to them that in order to give a concept-analysis of some noun phrase like *a man who is tall*, there needs to be a sentential analysis of the embedded relative clause (as well as an analysis of the contribution of ‘a’, as mentioned above in Sect. 8.1.2).

This puzzle is at least as old as Plato, and has provoked considerable discussion over the centuries. One type of answer might be to postulate a copula-like intermediary, call it ‘exemplification’, with the idea that it is this intermediary that causes the other items to merge into a full sentence. But as one can pretty easily see, this just leads to some sort of infinite regress: for, now there needs to be a further explanation of how it is that ‘Kim’ (or the person corresponding to that name) can combine with this newly postulated object to then be able to combine with ‘runs’ (or the object corresponding to that predicate). This is known as Bradley’s Regress, or The Unity of the Proposition/Sentence problem.

Objectivists have proposed various answers to the Unity problem, one of the most popular being some variant on the metaphor of *saturated/unsaturated* from Frege (1892a). The idea is that singular terms are such that they are designed to designate “independent” or “saturated” objects, whereas predicate terms are designed to designate “unfilled” or “unsaturated” items, namely “concepts”. (Recall that for Frege and the Objectivists, a concept is a non-mental, abstract entity.) When these two different kinds of entities “meet”, the saturated object “fills in” the unsaturated object, producing a new, saturated object—a complete sentence (which expresses a “thought”, in Fregean terminology... where “thought” is also understood as a non-mental, abstract item).

Well, whatever one thinks of such metaphorical accounts (and something akin to this is adopted in Formal Semantics’ model-theoretic theories), it is at least some *sort* of explanation, showing that there might be some wiggle room in Objectivist theories that would provide an appropriate description of what is going on.

However, it seems that Subjectivist accounts are not in such a good place, at least not if they hold that lexical concepts designate mental items like concepts. For it seems that this would presume there to be a concept of INSTANTIATION—sort of a mental version of the linguistic *exemplification*, and now it appears that the Bradley regress can start. Subjectivists of the sort we have been considering will respond by claiming that *conceptual combination* is how they handle this topic. But it should be noted that conceptual combination in these works is defined for noun-noun and adjective-noun combinations (perhaps also for adverb-verb combinations?). But this won’t work for sentence generation: perhaps the conceptual combination of DOG and

TAIL will yield a concept that somehow contains dogs with tails, but it isn't a concept corresponding to *Dogs have tails*.

The Bradley Regress puzzle should encourage Subjectivists to engage in serious theory construction for such a further mental operation. But it has to be an *operation*, and not a concept. For, a concept will only reinstate the Bradley Regress. So, besides conceptual combination, and whatever other related operations are required to accommodate the types of other combination I mentioned in the previous subsection, it is absolutely necessary that there be some further type of operation that forms "complete mental thoughts" out of concepts. And what should these be called? It seems to me that it would be much too confusing to simply reuse the term 'concept', so I will start using 'thought'. This maintains the parallel of usages with the (Fregean) Objectivists: they use 'concept' to refer to an objectively existing, abstract object; they use 'thought' to refer to an objectively existing, abstract object that can be evaluated in a truth-valued manner (or other manner for different speech acts than assertion). Subjectivists in the proposed usage employ 'concept' to refer to subjective items of individual minds that correspond to items of the non-sentential syntactic categories; they use 'thought' to refer to the mental complex that can be evaluated as a complete judgment.

8.3 *REAL Negation, Conjunction, Conditional, and Disjunction*

We've seen that English easily allows noun phrases, adjectival phrases, adverbial phrases, and verb phrases all to be conjunctive and disjunctive. English also allows for negative verb phrases, adjectival phrases, and adverbial phrases. And there is also some opening for conditional phrases of these various linguistic sorts. But this is not—at least, is not according to most philosophical and formal semantic theories—the major use of conjunction, disjunction, and negation. For that use, say the Objectivists, we negate, conjoin or disjoin sentences.⁵¹ But there is no provision within the Subjectivist conception of a mental regime of concepts for this type of combination. (For one thing, there has not been any explanation of how to generate thoughts from simpler concepts.)

But suppose we grant the Subjectivists that they can generate thoughts—that is, that they have some sort of concept combination that will generate judgments rather than just merged noun-like concepts. And further suppose they have some operations to generate concepts that are negations, conjunctions, and disjunctions of noun-, verb-, adjective-, and adverb-concepts. They still haven't done what most formal semanticists see as the central point of these notions: the ability to form thoughts which correspond to complex statements. The sort of work that is in the literature seems aimed at somewhat different targets, such as *describing* the various effects that different of these constructions have in specific cases as opposed to large-scale

⁵¹Or propositions, if one is in the market for using abstract entities rather than mental concepts.

rules that would apply universally. Those problems seem simply not to be the ones that are addressed, and so it is not so surprising that the investigations don't yield such overarching theoretical proposals.

Instead there are "local" accounts. But note, for the present case under discussion, that negation, conjunction, disjunction, and conditional cannot be concepts, under the pain of running again into the Bradley regress. But also it is not so easy to deal with these complex sentential operations as it is to deal with the corresponding ones operating on concepts. One example is discussed in Pelletier (1998), where the idea was that negation might be an operation that transforms "regular" concepts into some other type of concept. For example, the conceptual operation of negation might transform a verb-concept like IS TALL into a disjunctive concept like IS EITHER SHORT OR MIDDLE-HEIGHT. Or, it might transform the verb-concept IS RED into IS SOME COLOUR THAT IS CONTRARY TO RED, which would presuppose that colour concepts (and maybe all concepts generally) fall into "incompatibility classes", where nothing can manifest any two different concepts in that incompatibility class. And there are other possibilities.

This has the effect of making all sentence negations into predicate negations, and it is far from obvious that this sort of move can really be maintained. It is also far from obvious whether there is anything similar that can be done with the other "thought-combining concepts".

But even if there were some such way to endorse all the earlier additions to the mental world of concepts, and even if there were some mental analogue to the Objectivist notion of "saturation" as a way to solve the Unity of the Thought issue, it seems that the Subjectivists would just be mirroring the Objectivist account, but with a very complex and otherwise unmotivated set of operations. And then, such rumination makes one want to ask: Why bother? Even if it could all be done, would there have been any advance at all in explaining how natural language describes the world? Or even what is possibly more to the point from the Subjectivist point of view: has there been any advance at all in explaining how people use natural language to describe the world? It is hard to see what it might be.

It seems to me that the only reasonable answer from the Subjectivist camp to such rumination would have to emphasize the role of concepts in a *Belief+Desire*⇒*Action* framework. It is in this realm that we can see people making use of concepts that do not correctly mirror reality. ("She eats brown eggs rather than white eggs because she thinks they have less cholesterol"). It seems that an appropriate account of this subject's concepts would somehow explain such an objectively strange behaviour by using some such account of his/her beliefs and desires.) From the Objectivist camp, presumably the only appropriate answer to such rumination would be to show how "objective truth" can alter conceptual connections. And I think it is in this interaction of the two realms that the correct account of meaning lies, including why both are required.

9 (Some) Concluding Remarks

I have tried in this paper to lay out the (sometimes) subtle interactions that are inherent in a number of distinct commitments held by many theorists who study the semantics of language. A major division among these theorists is between those who think to study language by an investigation of how language describes or attaches to “the world” versus those who think of it primarily in terms of how it describes “mental structures” of users of the language. This is what I called the Objectivist versus Subjectivist distinction. This difference in viewpoint impinges on several topics of interest. A simple but confusing topic concerns the difference in intent when the two groups employ terms like ‘concept’ and ‘thought’, where Objectivists, following Frege, think of these as “public”, objective entities that can be “grasped” by many different language users, while Subjectivists think of them as “private” and existing only in the mind of a single language user. (Hence, any talk of a shared concept then becomes an issue of the similarity between the different concepts and thoughts possessed by the different speakers.)

Although it might appear on the surface that this is the entire difference between the philosophers-formal semanticists (the Objectivists) and the cognitive linguists-psychologists (the Subjectivists), there are further ramifications that may not be so obvious from just that simple difference. For, it is this that makes the Subjectivists think that meanings are not compositional, because of data about such topics as conceptual combination (for the cognitive psychologists) and the choice of meaning specificity by a speaker and hearer (for the cognitive linguists). And on the other hand, the separation of a speaker’s mental life from “meaning in the world” makes the philosopher/semanticist group able to hold on to semantic compositionality even in the face of the psychological data. They say that since mental concepts are not compositional, that is yet further evidence (besides what they believe they have shown) that such concepts can’t be meanings—since semantic compositionality is true.

This difference furthermore opens the door to distinct accounts of how meanings are to be specified. In the Objectivist version of this, meanings are organized in what I called an Atomistic manner. There are smallest units of meaning (words, or whatever) and these smallest units combine in one or another way to form the meaning of larger and larger units. The dispute over compositionality within the Objectivist camp then becomes whether this combination follows a notion of Functional Compositionality, or whether there can be exceptions. Note that this dispute is entirely about “meaning in terms of the objective world” and has no direct reflex into the subjective contents of a person’s mind.

But when it comes to the Subjectivist side of things, issues of compositionality can take two different forms. One is a kind of mental “mirror” of the Objectivist viewpoint, where there are smallest units that bear meaning—lexical items, let’s say. But these lexical items mean concepts and these concepts have a very rich structure, replete with “encyclopedic facts”. Nonetheless, the notion of Functional Compositionality *could* still apply to this picture, if only the meanings of more complex phrases could be described in terms of a function applied to the very complex mean-

ings of the simpler parts. As I mentioned, most researchers think this is not possible, although in my mind the jury is still out on this until the Subjectivist theory of complex concepts is filled out more fully. As it is currently given, the majority of research concerns the concept-combination of two particularly simple noun-based concepts. And for them it does indeed seem to be difficult to give a compositional treatment—although one should always keep in mind that functional compositionality is a particularly powerful tool with many unexpected strengths to capture under its aegis all sorts of phenomena in the objective realm. (Many of these are recounted in Pelletier 2016.) Perhaps the same strengths can be employed in the mental realm, especially when the “meaning” of concepts is explicated in an atomistic way, as outlined in Sect. 6.4. But it is difficult to know in the absence of a more developed theory of complex concepts.

The other way that a Subjectivist theory of meaning might go is to say that a part of each concept’s meaning is all the liaisons that concept has with other concepts. This is different from the other form of complex concepts, because here each concept has every other related concept as a part of its meaning. And since this holds of these related concepts also, it follows that what is related to them becomes a part of the original concept’s meaning. As one can see, this makes conceptual meaning be a “network” or “web”, where every concept involves every other concept, and no concept has any meaning of its own.

The former of the two Subjectivist theories is—like the Objectivist theories—Atomistic in nature. While the latter theory is Wholistic, leading to a holistic theory of meaning. To my mind, there are a number of devastating objections to any version of this latter theory. And so I rejected this version of Subjectivism, but not the other version, which is the general sort of account that I see being plausible for one half of a two-tier semantic theory.

It is not surprising, though, that Subjectivists of both types claim that semantic compositionality is false. Holistic semantic theories are, by definition, not Atomistic, and hence stand no chance of being compositional. The other type of Subjectivism, where concepts corresponding to lexical items contain a wealth of encyclopedic information, could be compositional if these presumed complex meanings happened to be organized appropriately. But at least so far as that idea has been developed by this point in time, they do not appear to be compositional. For the most part, the theoretical apparatus used for determining the meaning of complex mental items is so underspecified that it is in fact unable to create such complex items, much less create them in a way that would obey the strictures of semantic compositionality. But on the other hand, the Objectivists have not done two related things. They have not said how the lexical items that are at the terminal nodes of a syntactic tree are assigned their meanings other than to say that these meanings are “from the world”; and they have not described what relationship there is between this “meaning in the world” and a language-user’s mental life or concepts. (Neither for the lexical items nor for larger linguistic units like sentences.)

Do the Objectivists need to give such descriptions? After all, they might say, the enterprise of formal semantics is pretty much defined as articulating how sentences encode information about the world, and it has no particular place for human menta-

tion, except perhaps in the case of finding semantic treatments for attitude reports.⁵² And the task of a formal semantics is to explain how complex meanings are constructed from simple meanings. For that purpose one needs to start with the simple meanings as a given.

Although both of those claims are true, at least as the task is conceived by the Objectivists, they seem also to have forgotten their rationales for finding semantic compositionality attractive in the first place: the arguments from understandability, from productivity, from learnability, and from creativity. Certainly all of these arguments make the claim that the justification of semantic compositionality lies in the way it facilitates mental processing. And thus, even if the picture that “language is about the world” is correct, these arguments nevertheless presuppose or require that language also embody features of one’s mental life.

A natural place for the Objectivists to find this missing connection is to claim that the terminal nodes of their meaning structures are given by mental items, such as concepts. And then they can go ahead and—from this starting point—be entirely compositional. In fact, I’m sure that most of the philosophers-semanticists who give any thought to the issue probably believe something along those lines. But let’s look at what this would entail.

If the psychological results are correct—and they have been replicated over and over, at least in regards to simple binary combination—that mental concepts allow both for differences in degree of typicality of the concept and also for differences in degree of membership (Hampton and Jönsson 2012, pp. 386–87), then this would have to be a feature of the meanings at the terminal nodes of the semantic structure. The lexical item ‘furniture’, for instance, would have to admit *in the world* of different degrees of truth (and not merely different degrees of typicality) for ‘This chair is furniture’ and ‘This clock is furniture’. I think philosophers-semanticists would not countenance this.⁵³

In response to all these competing pressures, I surveyed the prospects for a two-tiered semantic theory: one aspect being the Objectivist’s “connection to the world” and the other aspect being the Subjectivist’s account of mental activity. As it turns out, some of the works that claim to provide a two-tier theory actually are within just one of the Subjectivist or Objectivist camp, in that both of their tiers have their foundational starting point, or common nexus, entirely within one camp. I also mentioned a few theories that appear to be genuinely two-tiered, but noted that they require further input, especially from the Subjectivist side of the theory. To agree to anything like the two-tiered theory I outlined, one needs to accept two claims: First, that linguistic utterances are to be judged as to how well they “fit” the (external) world. One way to put this, and I did, is to say that *what is meant*^o by an utterance concerns the actual world, and not any conceptual representation. The two come apart when

⁵²Such as *John wants to marry a princess*, the ambiguity of which (a specific princess versus any princess) seems to invite an analysis that would involve John’s mental states. And also, since human mentation is actually occurring in the physical world, it must find a proper place even in Objective semantics as something to be described.

⁵³I’m still speaking generically!.

the representation is incorrect in one way or another, and evaluation—whether the utterance is true or is false (in the case of declaratives) or otherwise appropriately felicitous (in the case of other speech acts)—is provided by the fit with the world. And secondly, that the appropriate explanation of human actions of all sorts of types is due to the meaning^s that an agent uses to *conceptualize* the world: how its beliefs and desires lead to actions that are in accordance with those beliefs and desires, and not to how the beliefs and desires match up with the way the world (really) is.

So the upshot of this casual survey of work from linguists, philosophers of language, and cognitive psychologists seems to be twofold. First, that formal semanticists and philosophers of language make theories that have “meaning” be a feature of how language “fits the world”—usually some sort of intensionalized, possible-world-oriented truth-theory, or (in other hands) an information-theoretic dynamic framework. In doing this they have no obvious place to attach this theory to “meaning” as some feature about how a speaker-hearer dynamic can possibly work. This seems to be a very undesirable state of affairs for a theory of meaning.

On the second hand, the cognitive linguists and the cognitive psychologists seem to have left out the way language is intended to “describe the world”, as they emphasize instead the way that an individual’s mental concepts work in classifying other aspects of their mental life. The attempt to make it be grounded, embodied, situated, or enactive appear to lack convincing detail. In particular, it is not so clear that any expedient from this general point of view *can* succeed, at least according to the Objectivists, since it continues to employ mental “information” (sensations, beliefs about the speech situation, and impressions about ones intentions). It will therefore always be “within the mental world” and not touch “reality”.

Only a two-tier theory can do both jobs.

Acknowledgements Thanks to Greg Carlson for discussions and for comments on an earlier draft. Thanks also to the editors of this volume, James Hampton and Yoad Winter, for their searching critiques of a slightly later draft. Their comments revealed numerous shortcomings in that version, but unfortunately, I have not been able to respond to them all. However, I hope to have ameliorated at least some of their concerns! Some of the alterations came about by discussions with my polymath colleague, Allen Hazen.

References

- Alxatib, S., & Pelletier, F. J. (2011). The psychology of vagueness: Borderline cases and contradictions. *Mind and Language*, 26, 287–326.
- Armstrong, S., Gleitman, L., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, 13, 263–308.
- Asher, N. (2011). *Lexical meaning in context: A web of words*. Cambridge: Cambridge UP.
- Atkins, S., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford: Oxford UP.
- Bar-Hillel, Y., & Eifermann, R. (1970). Who is afraid of disjunctive concepts? A case study in the genesis of pseudo-problems. *Foundations of Language*, 6, 463–472.
- Barsalou, L. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.

- Barsalou, L. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science*, 2, 716–724.
- Barsalou, L. (2017). Issues for psychologically plausible theories of conceptual combination. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 223–244). London: Springer.
- Boodin, J. (1939). *The social mind: Foundations of social philosophy*. New York: Macmillan Co.
- Bourne, L. (1982). Typicality effects in logically defined categories. *Memory and Cognition*, 10, 3–9.
- Brandom, R. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Cambridge, MA: Harvard UP.
- Brandom, R. (2000). *Articulating reasons: An introduction to inferentialism*. Cambridge, MA: Harvard UP.
- Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, 4, 73–121.
- Burling, R. (2005). *The talking ape: How language evolved*. Oxford: Oxford University Press.
- Bybee, J. (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication*, 2, 249–278.
- Bybee, J., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, 22, 381–410.
- Carlson, G. (2010). Generics and concepts. In F. Pelletier (Ed.), *Kinds, things, and stuff* (pp. 16–35). New York: Oxford UP.
- Casati, R., & Varzi, A. (2015). Events. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/win2015/entries/events/>.
- Cheney, D. L., & Seyfarth, R. M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review*, 22, 135–159.
- Chierchia, G., & McConnell-Ginet, S. (2000). *Meaning and grammar: An introduction to semantics* (2nd ed.). Cambridge, MA: MIT Press.
- Chomsky, N. (1977). *Essays on form and interpretation*. New York, NY: North-Holland.
- Chomsky, N. (1992). Explaining language use. *Philosophical Topics*, 20, 205–231.
- Chomsky, N. (1995). Language and nature. *Mind*, 104, 1–61.
- Conant, M., & Trabasso, T. (1964). Conjunctive and disjunctive concept formation under equal-information conditions. *Journal of Experimental Psychology*, 67, 250–255.
- Connell, L., & Lynott, D. (2014). Principles of representation: Why you can't represent the same concept twice. *Topics in Cognitive Science*, 6, 390–406.
- Croft, W. (2001). *Radical construction grammar*. Oxford: Oxford UP.
- Dirven, R., & Fried, V. (Eds.). (1987). *Functionalism in linguistics*. Amsterdam: John Benjamins.
- Dummett, M. (1991). *The logical basis of metaphysics*. London: Duckworth.
- Evans, V. (2009). *How words mean: Lexical concepts, cognitive models, and meaning construction*. Oxford: Oxford UP.
- Evans, V. (2015). *The crucible of language: How language and mind create meaning*. Cambridge, UK: Cambridge UP.
- Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Fodor, J. (1980). Methodological solipsism considered as a research strategy in cognitive science. *Behavioral and Brain Sciences*, 3, 63–73.
- Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Fodor, J. (1998). *Concepts: where cognitive science went wrong*. Cambridge, MA: MIT Press.
- Fodor, J. (2000). *The mind doesn't work that way: The scope and limits of computational psychology*. Cambridge, MA: MIT Press.
- Fodor, J. (2001). Language, thought, and compositionality. *Mind and Language*, 16, 1–15.
- Fodor, J., & Lepore, E. (1991). Why meaning (probably) isn't conceptual role. *Mind and Language*, 6, 328–343.
- Fodor, J., & Pylyshyn, Z. (2015). *Concepts without meaning*. Cambridge, MA: MIT Press.

- Francez, N. (2014). Proof-theoretic semantics for contextual domain restriction. *Journal of Language Modelling*, 2, 249–283.
- Francez, N. (2015). *Proof-theoretic semantics*. London: College Publications.
- Francez, N., & Ben-Avi, G. (2015). A proof-theoretic reconstruction of generalized quantifiers. *Journal of Semantics*, 32, 313–371.
- Francez, N., & Dyckhoff, R. (2010). Proof-theoretic semantics for a natural language fragment. *Linguistics and Philosophy*, 33, 447–477.
- Francez, N., Dyckhoff, R., & Ben-Avi, G. (2010). Proof-theoretic semantics for a sub-sentential phrases. *Studia Logica*, 94, 381–401.
- Frege, G. (1891, January). Funktion und Begriff. Vortrag, gehalten in der Sitzung der Jenischen Gesellschaft für Medizin und Naturwissenschaft. Translated as “Function and Concept”.
- Frege, G. (1892a). Über Begriff und Gegenstand. *Vierteljahresschrift für wissenschaftliche Philosophie* 16, 192–205. Translated as “On Concept and Object”.
- Frege, G. (1892b). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100, 25–60. Translated as “On Sense and Reference” (or “On Sense and Meaning”, by different translators).
- Fried, M., & Östman, J.-O. (Eds.). (2004). *Construction grammar in a cross-language perspective*. Amsterdam: John Benjamins.
- Furth, H. (1963). Classification transfer with disjunctive concepts as a function of verbal training and set. *Journal of Psychology*, 55, 477–485.
- Gaeraerts, D., & Cuyckens, H. (2007). Introducing cognitive linguistics. In D. Gaeraerts & H. Cuyckens (Eds.), *Oxford handbook of cognitive linguistics* (pp. 1–21). Oxford: Oxford UP.
- Gagné, C. (2002). Lexical and relational influences on the processing of novel compounds. *Brain and Language*, 81, 723–735.
- Gagné, C., & Spalding, T. (2006). Conceptual combination: Implications for the mental lexicon. In G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 145–168). New York: Oxford UP.
- Gagné, C., Spalding, T., & Kostecky, M. (2016). Conceptual combination, property inclusion, and the Aristotelian-Thomistic view of concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and Concepts in Linguistics and Psychology* (pp. 245–268). London: Springer.
- Goddard, C. (2002). The search for the shared semantic core of all languages. In *Goddard and Wierzbicka* (Vol 1, pp. 5–40).
- Goddard, C., & Wierzbicka, A. (1994). *Semantic and lexical universals*. Philadelphia: Benjamins.
- Goddard, C., & Wierzbicka, A. (2002). *Meaning and universal grammar: Theory and empirical findings* (Vols. 1 & 2). Philadelphia: John Benjamins.
- Goldberg, A. (1995). *Constructions: A Construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Greenhill, S. J., Atkinson, Q. D., Meade, A., & Gray, R. D. (2010). The shape and tempo of language evolution. *Proceedings of the Royal Society, B*, 277, 2443–2450.
- Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard UP. Contains Grice’s articles on language, and his later thoughts about his earlier works.
- Halliday, M. (1978). *Language as social semiotic: The social interpretation of language and meaning*. University Park, MD: University of Maryland Press.
- Hampton, J. A. (1979). Polymorphic concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 18, 441–461.
- Hampton, J. A. (1982). A demonstration of intransitivity in natural categories. *Cognition*, 12, 151–164.
- Hampton, J. A. (1988a). Disjunction of natural concepts. *Memory and Cognition*, 16, 579–591.
- Hampton, J. A. (1988b). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology*, 14, 12–32.
- Hampton, J. A. (1997). Conceptual combination: Conjunction and negation of natural concepts. *Memory and Cognition*, 25, 888–909.

- Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science*, 31, 355–384.
- Hampton, J. A. (2011a). Concepts and natural language. In R. Belohlavek & G. Klir (Eds.), *Concepts and fuzzy logic* (pp. 233–258). Cambridge, MA: MIT Press.
- Hampton, J. A. (2011b). Conceptual combinations and fuzzy logic. In R. Belohlavek & G. Klir (Eds.), *Concepts and fuzzy logic* (pp. 209–232). Cambridge, MA: MIT Press.
- Hampton, J. A. (2017). Compositionality and concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 95–122). London: Springer.
- Hampton, J. A., & Jönsson, M. (2012). Typicality and compositionally: The logic of combining vague concepts. In M. Werning, W. Hintzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 385–402). Oxford: Oxford UP.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.
- Hinzen, W., & Sheehan, M. (2013). *The philosophy of universal grammar*. Oxford: Oxford UP.
- Hurfurd, J. R. (2000). Social transmission favours linguistic generalization. In C. Knight, M. Studdert-Kennedy, & J. R. Hurfurd (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 219–230). Cambridge: Cambridge University Press.
- Johnson, M. (2014). Compositionality. Internet Encyclopedia of Philosophy. <http://www.iep.utm.edu/composit>.
- Jurafsky, D., & Martin, J. (2008). *Speech and language processing* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–566). Oxford: Oxford UP.
- Kay, P. (1995). Construction grammar. In J. Verschueren, J.-O. Östman, & J. Blommaert (Eds.), *Handbook of pragmatics* (pp. 171–177). Amsterdam: John Benjamins.
- Komatsu, L. (1992). Recent views of conceptual structure. *Psychological Bulletin*, 112, 500–526.
- Krifka, M., Pelletier, F. J., Carlson, G., ter Meulen, A., Chierchia, G., & Link, G. (1995). Genericity: An introduction. In G. Carlson & F. J. Pelletier (Eds.), *The generic book* (pp. 1–124). Chicago: University of Chicago Press.
- Kripke, S. (1980). *Naming and necessity*. Cambridge MA: Harvard UP.
- Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics. 2nd edition Cambridge University Press, 2006.
- Labov, W. (1975). Empirical foundations of linguistic theory. In R. Austerlitz (Ed.), *The scope of American linguistics* (pp. 77–133). Lisse: Peter de Ridder.
- Langacker, R. (1987). *Foundations of cognitive grammar. Vol. 1: Theoretical prerequisites*. Stanford: Stanford UP. Chapter 12: Composition.
- Langacker, R. (1990). *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.
- Langacker, R. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford UP.
- Laurence, S., & Margolis, E. (1999). Concepts and cognitive science. In S. Laurence & E. Margolis (Eds.), *Concepts: Core readings* (pp. 3–82). Cambridge, MA: MIT Press.
- Leben, D. (2015). Neoclassical concepts. *Mind and Language*, 30, 44–69.
- Lewis, D. (1970). General semantics. *Synthese*, 22, 18–67.
- Li, C. (2002). Missing links, issues and hypotheses in the evolutionary origin of language. In T. Givón & B. Malle (Eds.), *The evolution of language out of pre-language* (pp. 83–106). Amsterdam: John Benjamins.
- Locke, J. (1690). *Essay concerning human understanding*. Oxford: Clarendon Press. Edited, with forward, by Peter Niddich; published 1975.
- Machery, E. (2009). *Doing without concepts*. Oxford: Oxford UP.
- MacWhinney, B. (2005). The emergence of grammar from perspective taking. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 198–223). Cambridge: Cambridge University Press.

- Manning, C., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- Martin-Löf, P. (1980). *Intuitionistic type theory*. Naples: Bibliopolis.
- McGilvray, J. (1998). Meanings are syntactically individuated and found in the head. *Mind and Language*, 13, 225–280.
- McNally, L., & Boleda, G. (2017). Conceptual vs. referential affordance in concept composition. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 245–268). London: Springer.
- Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Newstead, S., & Griggs, R. (1983). The language and thought of disjunction. In J. S. Evans (Ed.), *Thinking and reasoning: psychological approaches* (pp. 76–106). London: Routledge & Kegan Paul. Reprinted 2013 by Taylor & Francis, in their Psychology Revivals series.
- O’Grady, W. (2008). The emergentist program. *Lingua*, 118, 447–464.
- Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- Osherson, D., & Smith, E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 11, 35–58.
- Östman, J.-O., & Fried, M. (Eds.) (2004). *Construction grammars*. John Benjamins.
- Pagin, P. (2009). Compositionality, understanding, and proofs. *Mind*, 118, 713–737.
- Pagin, P. (2012). Communication and the Complexity of Semantics. In M. Werning, W. Hinzen & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 510–529). Oxford: Oxford UP
- Parsons, T. (1990). *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.
- Partee, B. (1979). Semantics—mathematics or psychology? In R. Bäuerle, U. Egli, & A. von Stechow (Eds.), *Semantics from different points of view* (pp. 1–14). Berlin: Springer.
- Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge: Cambridge University Press.
- Pelletier, F. J. (1998). Thinking of ‘not’. In X. Arrazola, K. Kortá, & F. J. Pelletier (Eds.), *Discourse, interaction, and communication* (pp. 37–47). Dordrecht: Kluwer.
- Pelletier, F. J. (2012). Holism and compositionality. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 149–174). Oxford: Oxford UP.
- Pelletier, F. J. (2013). Context, compositionality, and brevity. In L. Goldstein (Ed.), *Brevity* (pp. 178–197). Oxford: Oxford UP.
- Pelletier, F. J. (2016). Semantic compositionality. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics*. Oxford: Oxford UP.
- Pelletier, F. J., Elio, R., & Hanson, P. (2008). Is logic all in our heads? From naturalism to psychologism. *Studia Logica*, 88, 3–66.
- Power, C. (1998). ‘Old wives’ tales’: The gossip hypothesis and the reliability of cheap signals. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language* (pp. 111–129). Cambridge: Cambridge University Press.
- Prawitz, D. (2006). Meaning approached via proofs. *Synthese*, 148, 507–524.
- Prinz, J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. Cambridge, MA: MIT Press.
- Putnam, H. (1975). The meaning of ‘meaning’. In K. Gunderson (Ed.), *Mind, language and reality* (pp. 215–271). Minneapolis: University of Minnesota Press.
- Putnam, H. (1981). Brains in a vat. In H. Putnam (Ed.), *Reason, truth and history* (pp. 1–21). Cambridge, UK: Cambridge UP.
- Recanati, F. (2012). Compositionality, flexibility, and context dependence. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 175–191). Oxford: Oxford UP.
- Rohrer, T. (2007). Embodiment and experimentalism. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 25–47). Oxford: Oxford UP.

- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–232.
- Sainsbury, M. (2001). Two ways to smoke a cigarette. *Ratio*, 14, 386–406.
- Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 3, 532–631.
- Schein, B. (1993). *Plurals and events*. Cambridge, MA: MIT Press.
- Schmid, H.-J. (2007). Entrenchment, salience, and basic levels. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 117–138). Oxford: Oxford UP.
- Scholz, B. C., Pelletier, F. J., & Pullum, G. K. (2016). Philosophy of linguistics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/win2016/entries/linguistics/>.
- Schroeder-Heister, P. (2016). Proof-theoretic semantics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/win2016/entries/proof-theoretic-semantics/>.
- Schubert, L. (2015). Computational linguistics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/spr2015/entries/computational-linguistics/>.
- Seggie, J. (1969). Levels of learning involved in conjunctive and disjunctive concepts. *Australian Journal of Psychology*, 21, 325–333.
- Shapiro, L. (2014). *The Routledge handbook of embodied cognition*. Oxford: Routledge.
- Sinha, C. (2007). Cognitive linguistics, psychology, and cognitive science. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 1266–1294). Oxford: Oxford UP.
- Skinner, B. (1957). *Verbal behavior*. Acton, MA: Copley Publishing.
- Smith, E., & Medin, D. (1981). *Categories and concepts*. Cambridge, MA: Harvard UP.
- Smith, E., & Osherson, D. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, 8, 337–361.
- Smith, E., Osherson, D., Rips, L., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, 12, 485–527.
- Smith, E., Shone, E., & Rips, L. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 81, 214–241.
- Smuts, J. (1926). *Holism and evolution*. London: MacMillan.
- Spalding, T., & Gagné, C. (2008). CARIN theory reanalysis reanalyzed: A comment on Maguire, Devereux, Costello, and Cater. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1573–1578.
- Szabó, Z. (2012). The case for compositionality. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 64–80). Oxford: Oxford UP.
- Taylor, J. (2007). Cognitive linguistics and autonomous linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 566–588). Oxford: Oxford UP.
- Tomasello, M. (1998). Introduction. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure*. Mahwah, NJ: Lawrence Erlbaum.
- Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology*, 2, 239–289.
- Turner, M. (1991). *Reading minds: The study of English in the age of cognitive science*. Princeton, NJ: Princeton UP.
- Tye, M. (2009). A new look at the speckled hen. *Analysis*, 69, 258–263.
- Van Valin, R. (2003). Functional linguistics. In M. Aronoff & J. Rees-Miller (Eds.), *The handbook of linguistics* (pp. 319–366). Oxford: Blackwell.
- von Humboldt, W. (2000). Character of languages; poetry and prose. In M. Losonsky (Ed.), *Humboldt on language* (pp. 148–181). Cambridge, UK: Cambridge UP.
- von Savigny, E. (1988). *The social foundations of meaning*. Berlin: Springer.
- Westerståhl, D. (2012). Compositionality in Kaplan-style semantics. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 192–219). Oxford: Oxford UP.

- Wierzbicka, A. (1996). *Semantics: Primes and universals*. Oxford: Oxford UP.
- Wildgen, W. (2008). Sketch of an evolutionary grammar based on comparative biolinguistics. In L. Röska-Hardy & E. Neumann-Held (Eds.), *Learning from Animals? Examining the nature of human uniqueness* (pp. 45–59). London: Psychology Press.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9, 219–248.
- Wilson, R. A., & Foglia, L. (2016). Embodied cognition. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/win2016/entries/embodied-cognition/>.
- Winter, Y. (2017). Critical typicality: Truth judgements and compositionality with plurals and other gradable concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 163–190). London: Springer.
- Wisniewski, E. (1997). When concepts combine. *Psychonomic Bulletin & Review*, 4, 167–183.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Basil Blackwell. Translated by G.E.M. Anscombe.
- Worden, R. (1998). The evolution of language from social intelligence. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language* (pp. 148–166). Cambridge: Cambridge University Press.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Compositionality and Concepts

James A. Hampton

Abstract In this chapter I aim to explain how psychology understands concepts, and why there is a need for semantic theory to take on the challenge of psychological data. All of the contributors to this volume are (presumably) in the business of trying to understand and explain how language has meaning, and the primary source of evidence for this has to be our intuitions of what things mean. Furthermore, if my semantic intuitions (as a theorist) are out of kilter with those of the common language user, then it is my theory which should be called into question and not the lay intuition. This chapter describes a range of results from my research program over the last 30 years, some old and some new, with the aim of giving a general account of using Prototype Theory as a way to explain semantic intuitions.

1 Concepts and Prototypes

In *The Compositionality Papers*, Fodor and Lepore (2002) return frequently to a “knock-down” argument against the suggestion that concepts might be prototypes. Concepts, they argue, must be compositional. It must be possible, if one accepts the representational theory of mind, to explain how the meaning of a complex phrase is based solely on the meaning of the elements from which it is constructed, plus the syntactic structure into which they are placed. To account for our ability to understand the meaning of sentences such as (1) and (2) and countless other similar sentences the semantic system needs a set of fixed symbols to represent the conceptual atoms in the sentences (John, Mary, Bill, loves, hates) which can then be inserted into suitable syntactically structured sentence frames to yield the appropriate meaning for the sentence as a whole.

J.A. Hampton (✉)
Department of Psychology, City, University of London, London, UK
e-mail: j.a.hampton@city.ac.uk

© The Author(s) 2017
J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts*
in *Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_4

- (1) John loves Mary.
- (2) John loves Mary, but Mary loves Bill and Mary hates John.

They claim that without this type of compositionality it is not possible to provide an account of how thoughts (and indeed utterances) can express ideas.

Given the claim that concepts must be compositional, the argument continues that prototypes are in fact anything but compositional. For example, the prototypical pet fish is not simply the prototype pet conjoined with the prototype fish. Indeed something like a goldfish or guppy, while being a good match for the pet fish prototype, has little in common with the cats and dogs that are typical pets, or the cod and trout that are typical fish (Osherson and Smith 1981).

More generally, the prototype of any particular complex noun phrase (should it have one) will not be derivable from the prototypes of the content words of which it is composed. Hence concepts (which must be compositional) cannot be prototypes (which are often non-compositional).

The debate on compositionality has generated a large literature in semantics and philosophy—see for example the substantial collection of papers edited by Werning et al. (2012) For critical accounts on whether language is in fact compositional, particularly in respect of the systematicity of its grammar, see Johnson (2004) and Pullum and Scholz (2007). As Pelletier (2017) remarks, the issue probably represents one of the most substantial points of disagreement within cognitive science between the more empirically minded psychologists and linguists who study concepts and word meanings “in the mind”, language in use, and lay people’s semantic intuitions, and the more theoretically minded semanticists and philosophers whose interests cover broader issues of concepts as constituting lexical meanings within a particular language, and issues of sentential truth and logic. (Like Pelletier, I am speaking generically here).

As an empirical researcher into word meanings, I find myself strongly drawn to the conclusion that lexical meanings do not compose according to the rules of set logic applied to extensions. That is to say that the meaning of a complex phrase will not always be determined simply by the meanings of its components and their mode of combination. I will argue that the construction of complex concepts proceeds (most naturally) through the interactive combination of the intensional meanings of the individual concepts. Wherever two concepts are combined whose intensional contents overlap or interact semantically, then extensional compositionality of meaning will tend to fail. Moreover, if we consider thought rather than language, our capacities to combine concepts greatly exceed the simple combinatorial rules provided by extensional logic, as will become apparent in the last part of this chapter.

1.1 *Combining Prototypes*

A key piece of evidence for proposing that the category membership of complex concepts depends on more than their components and their mode of combination comes from a series of experiments that I and others conducted in the 1980s and 1990s. To illustrate, in Hampton (1988b, Experiments 2 and 3) I looked at how people interpret the phrases “Sports which are Games”, or “Games which are Sports”. These phrases were chosen because (at least at first sight) a standard semantic analysis would see the meaning as being a conjunction—that is the extension of both phrases should correspond quite simply to the intersection of the set of Games and the set of Sports. The two alternative phrasings come at the intersection in different ways (either finding the subset of sports that are also games, or vice versa) but the result should be the same.

To test this standard proposal, I provided 36 respondents with a list of 43 recreational activities selected to fall in all four possible combinations of being Sports or not, and being Games or not. Their first task was to judge whether each item was a Sport (together with a rating of degree of typicality or relatedness) and then to judge whether each item was a Game. (Order of tasks was balanced over subjects). Four weeks later the same individuals returned and decided which items were in one or other of the conjunctions mentioned above.

Based purely on their true/false judgments of “whether the general category name can be applied to a particular example”, the data were analysed to see whether in fact people followed an intersective rule, only saying Yes to the conjunction for those items (like Tennis or Football) where they had said Yes to both conjuncts. The results were very clear. People did not follow this rule. In fact 25% of games that were not sports, and 54% of sports that were not games, according to first week responses, were nevertheless categorized in the conjunction on week 4. This overextension was not attributable to a contrast effect between the categories, nor to randomness in people’s choices, since the equivalent inconsistent pattern of saying No to an item that had been judged as in both sets was much less frequent (11%). A further study showed that the effect was not driven by a response bias, since adding in new items to the list for the second phase so that the expected rate of Yes responses for the conjunction would be 50% (rather than 25%) had no influence on the rate of overextension of the other items.

A final experiment in this paper showed that the effect generalized to six other pairs of categories, including pets which are birds and dwellings which are buildings. Regression analysis was applied to predicting the mean degree of membership in the conjunction from means for each conjunct. Including an interaction term, on average 93% of the variance in mean membership ratings for the conjunction was explained. The analyses also showed that the relative clause form is non-commutative (for example, weapons which are tools are not equivalent to tools which are weapons) and that one of the two concepts often carries more weight in the prediction than the other (a phenomenon termed dominance).

Subsequent research has shown that this lack of respect for the conjunction rule is easily replicated across a range of domains and linguistic ways of expressing a conjunction (Storms et al. 1998; Hampton 1996; Jönsson 2015). It can be shown with the same individual making all three judgments, or with different individuals making each judgment. It also appears when the relative clause is negated. For example, a horse is not often considered to be a vehicle, but it is frequently considered to be a vehicle that is not a machine. Hampton (1988a) showed that the problem of non-extensional combination applies equally to disjunctions. A mushroom is never classed as a fruit, and only 50% of respondents called it a vegetable. However 90% decided that it was either a fruit or a vegetable.

The challenge of these results to extensional theories of semantics should be clear. These results do not just depend on judgments of typicality or even judgments of degrees of truth which might be subject to psychological biases (Osherson and Smith 1981, 1982). They simply reflect the common semantic intuitions of everyday language speakers about the applicability of complex phrases, and as such they require an explanation.

1.2 *Intensional Composition*

The account of overextension that I offered in Hampton (1987) was a relatively straightforward explanation based on the intensional properties of the concepts involved.

Classically speaking, when a complex concept is constructed as a conjunction, then the features that define it intensionally will be derived in a compositional fashion as a disjunction of the features of the two conjuncts. For example, to form the conjunction of people who are both singers and songwriters, one would look to set up a categorization procedure that would require potential candidates to have the union of the features of each set (person AND singer AND songwriter). To model the combination of categories such as Sports and Games, one can simply propose that Sports which are also Games should therefore be a composite prototype resulting from aggregating all the features commonly associated with either of the two conjuncts. As Sports they should be activities which involve exercise and training and as Games they should also involve competition and fun.

Most importantly, this way of modelling the construction of a conjunctive concept provides a neat explanation of why people are inconsistent in judging extensions of conjunctions. Consider the case where there are two prototype concepts A and B, each with 3 features. In line with Rosch's idea of family resemblance and prototypes, let us suppose that any item will belong in either concept if it has at least 2 of the 3 features of that concept.

Table 1 shows the resulting composite, in which the conjunction $A \wedge B$ is created with all six features in its prototype. Now if having 2 out of 3 features is sufficient to belong in either of the two concepts A or B, then it is natural to suggest that 4 out of 6 features should be sufficient to belong in their conjunction. Accordingly, Item 1

Table 1 A scheme for aggregating the attributes of two concepts A and B into a conjunctive concept $A \wedge B$. If a “two out of three” rule is used to determine category membership, then Item 1 will be a member of A, B and of the conjunction. However Item 2 will be a member of A and of the conjunction, but will not be a member of category B, having only one of the b attributes. The model predicts overextension of the conjunction through compensation (an excellent member of A can be in the conjunction $A \wedge B$ although only weakly connected to B)

Concept A
a1
a2
a3

Concept B
b1
b2
b3

Conjunction $A \wedge B$	Item 1	Item 2
a1	a1	a1
a2	a2	a2
a3	-	a3
b1	b1	b1
b2	b2	-
b3	-	-

which has 2 of each of the concept’s features should belong in the conjunction. However Item 2, which has 3 A features but only 1 B feature, passes the criterion for belonging in the conjunction, but fails to have enough features to be an example of Concept B. The result is that Item 2 would be overextended—it would belong in the conjunction but not in one of the conjuncts. Effectively, pooling together the features, even before any interaction between them has been considered (resulting in inheritance failure or the addition of new emergent features, see below), leads to the likelihood of inconsistent responding and overextension.

Hampton (1987) collected data on the attributes associated with each of the conjuncts and their conjunction for the same pairs of categories as were used in Hampton (1988b) described above, and traced the way in which attributes for the composite prototype are derived from the constituent components. The procedure involved two separate samples of participants. A first sample was divided into four groups. For a pair of concepts such as Sports and Games, group 1 listed attributes for Sports, group 2 for Games, group 3 for “Sports that are also Games” and group 4 for “Games that are also Sports”. They were asked to imagine that they had to define and describe the objects named to someone who was unfamiliar with them. They were to do this by listing on 10 blank lines the attributes or properties that were in any way involved in deciding if an object belongs in the named set. (Full details can be found in Hampton 1987, p. 58). Lists of around 30 attributes for each pair of concepts were then drawn up by including any attribute generated by at least 3 out of 10 participants in any of the four groups.

A second sample of participants was then used to judge the attributes for how important they were for defining each concept and each conjunction. The sample was divided into four groups exactly as before, but now each participant saw the list of attributes and made a rating judgment. “N” was to be chosen as a response if an

attribute was necessary for a category, responses “A”, “B” and “C” were used to indicate decreasing importance (A = very important, C = typically true but not very defining), “X” meant the attribute was not usually true, and “XX” meant the attribute was necessarily false of all possible examples of the concept. The results from this rating task were then used to assess how attributes of conjuncts are inherited by their conjunctions.

There were many interesting aspects to the results which I will not try to summarize here. As predicted, importance for a conjunct could be used to predict importance for a conjunction (Multiple R averaged around 0.8, close to the reliability of the scales at 0.85). The process is analogous to one of inherited traits, with traits possessed by both parent concepts being carried through to the offspring conjunction. Most notably there were some attributes which failed to be “inherited” by the composite, and others that “emerged” in the composite which were not in the constituents. Pets which are Birds, for example, lost some features of pets (cuddly) and some features of birds (fly south in winter), but gained other new features not seen in either concept alone such as (lives in cages) or (talks).

1.3 Prototypes as Intensions

I have argued that the way in which people interpret simple semantic rules such as relative clause modification can be accurately modelled by a deeper analysis of the intensional meaning of the words. What is the evidence then that such intensions have the prototype structure that leads to the patterns of overextension seen above?

If we consider most common content words in natural language—nouns, verbs, adjectives and adverbs—then it is often the case that neither the extension nor the intension are easy to pin down. The meaning of function words like prepositions is even harder. Consider for example the following common uses of “on” in English:

- (3) The cup is on the table.
- (4) I got paint on my shirt.
- (5) Harry is on holiday.
- (6) The train is on platform 2.

Any attempt to define the extension of situations to which “on” applies is likely to end up simply as a disjunctive list of different cases. The fact that prepositions do not easily translate between languages supports this claim (Bowerman and Choi 2003a, b). When I told a French friend that I had travelled to Paris “sur le train”, the puzzled response was to ask if it wasn’t very windy sitting on the roof. Similar problems arise with adjectives such as “fresh” or “open” (Murphy and Andrew 1993), with multiple inter-related senses determined by context. For further discussion see Rice (1992).

Returning to the (perhaps) simpler case of nouns, consider a simple everyday term such as “fish”. First there is a potential ambiguity arising from the domain of discourse. Fish features in cookery and food and as such its extension may include creatures such as squid, oysters and lobsters. Fish is also part of a commercial

industry, exploiting marine resources, and in the past the category extension included whales. Finally, fish may be taken to have a biological meaning. Unfortunately, those who pin their hopes on science to identify the “correct” extensional class, are due to be disappointed. Current scientific theory suggests there is no common ancestor to the different classes that we call fish. The term describes a disjunctive category with no role to play in biological theory. Thus, not only do we have these three different ways to place the term in context, but within each context the determination of what is in the extension becomes equally problematic. Should shell-fish be included in the culinary term—what about seahorses, rays or squid? The term is underspecified, as the psychological data clearly show (Hampton 1998; Hampton et al. 2006; McCloskey and Glucksberg 1978).

Since intensions are closely tied to extensions, it is not surprising to find that the intensional definitions of terms are equally difficult to pin down. Most encyclopaedia entries will state that all fish are cold-blooded. However, they then go on to say that some fish (like tuna) are not. It has proved very difficult to provide clear definitions for most of our vocabulary terms. There always seem to be exceptions.

Exceptions to the rule that there are always exceptions may be found when a concept has a particularly important role to play in the regulation of society. Then it will often be found to have an explicit definition. The definition of a “US dollar” or “British citizen” has a legal foundation which leads to a clear-cut differentiation into members and non-members of the class.

A slightly less clear example is provided by certain kinship terms, like “father” or “nephew” in English which are often found to be amenable to an analysis in terms of semantic components. As Goodenough (1965) describes it

A system of kin relationships rests on the established institutions and customs relating to membership in households, sexual rights, the definition of procreation, the legitimization of progeny as members of a jural community and the like.

In relation to “grandmother” Landau (1982) showed how both a definitional criterion (female having a grandchild) and a stereotypical age and appearance are seen in responses for both children and adults when selecting appropriate pictures.

Even when kinship terms such as “uncle” are extended to non-blood relations we are able to distinguish a “real” uncle from other kinds of uncle. On the other hand, the development of non-traditional families has led to the undermining of many kinship terms (see Lakoff’s 1987, discussion of “mother”), and terms such as “brothers” and “sisters” can be used with extended meaning to refer to others who share the speaker’s beliefs, goals or group membership.

Terms describing crimes (such as murder, theft, or fraud) are likewise provided with definitions by the legislature of each jurisdiction, so that juries can focus on establishing the facts of a case based on the evidence, rather than having to decide how to interpret the meaning of the words. (The latter task is left to judges in higher courts who aim to establish stable interpretations of the terms through reasoned argument about test cases and guessing the plausible intention of the law-makers). However, when a concept does not have this consequential weight resting upon it, it will usually resist easy definition.

The lack of a clear definition where one is needed can lead to expensive court cases, as the following extract from an article by Caroline Davies in the UK newspaper *The Guardian* of 27 April, 2015, shows:

Bridge, the genteel and physically unchallenging card game played by millions, may exercise the brain muscle, but is it a sport? That is the question taxing legal minds as a high court ruling on Monday paved the way for a courtroom battle to decide. The row centres on a refusal by Sport England to recognise the trick-taking game as a legitimate sport and thus eligible for lottery grants. The English Bridge Union claim it ought to be recognised as a “mind sport” and want Sport England’s refusal to do so declared unlawful.

Arguments presented to the court included the amount of physical activity involved (compared for example to rifle shooting), the health benefits of taking part, and the fact that other physical activities are not classified as sports. It was clear that lawyers on each side were seeking to find a plausible definition (intension) that would enable them to either include or exclude bridge from the category containing clear examples of sport such as tennis or football. (What is less clear is why the judge was willing to entertain the argument that the brain is a muscle!)

The issues involved here can be related to two fundamental issues in semantics—context sensitivity and vagueness. Perhaps the lack of clearly specified meanings of terms, and the consequent inconsistency in semantic intuitions results from the lack of a clear context. Alternatively, the difficulty of providing clear meanings may in fact result from those meanings being inherently unclear or vague, in the same way that scalar adjectives such as “tall” or “bald” have been shown to lack precision.

1.4 Context Sensitivity

The meanings of terms can change depending on context. Classic examples are scalar adjectives—a large ant is not as big as a small elephant. Alternatively, there is the example of fish described above, and the different contexts in which the term might be used.

But can context sensitivity fully explain the difficulty in defining extensions and intensions? In an attempt to find evidence for this suggestion, Hampton et al. (2006) ran a set of studies in which we manipulated the context in which people had to classify items in vague categories. We used eight different categories from different ontological domains, and created lists of 24 possible members in which we deliberately included clear cases, clear non-members of the category and about 12 cases that would be difficult to categorise. (The existence of borderline cases in natural categories was originally demonstrated by McCloskey and Glucksberg 1978, when they showed that not only were there many items showing substantial disagreement between people, but that people were also inconsistent in how they categorized those same items when returning a month later to do the task again.)

Our hypotheses for the study related to the idea that apparent vagueness in the category boundaries can be attributed to a lack of a clear context for the classification. We therefore had three main conditions in which we provided different contexts for the categorization. The first was a Neutral control condition which simply asked “Consider each of the following items and decide whether they belong in the category of _____”. A second condition, the Pragmatic condition, asked people to categorize items in categories “where people would expect to find them, so that they could be easily found”. Scenarios included an internet news group, a mail-order catalogue and a library index. Here we hoped to reduce vagueness because everyone would be attempting to mirror the behaviour of everyone else in the group. The third condition, the Technical condition, provided a set of contexts much like the case of Bridge and Sport described above. People were asked to imagine that they were advising a government agency controlling tax regimes (for Tools and Furniture), ecological reports (for Insects and Fish) or funding agencies (for Science and Sport). They were told that the classification would have important consequences and so they should try to classify “correctly”.

Participants worked through each list classifying items as Yes or No, and returned after 3–4 weeks to do the task again in the same condition as before.

Our prediction was that if lack of categorization context was contributing to vagueness, then various measures of vagueness would be reduced in the Technical and Pragmatic conditions. There should be better inter-subject agreement on classification, more stability in categorization decisions over time, a reduced correlation of categorization probability with simple ratings of Typicality, and a shift in the size of the categories, with Technical conditions yielding smaller categories. In the event, none of these predictions was generally supported by the data. Effectively categorization probability in all conditions was correlated at around 0.95 (the limit of measurement reliability) with judgments of item typicality in the category. A second study showed that requiring people to read the instructions aloud and reflect on them before starting to categorize had no effect on the results. There was no easily accessible “deeper” meaning for people to describe if asked to take the task more seriously. Finally, we looked at whether people would be less likely to give a “partial” or graded response in the Technical condition. In this last study, people were given a graded categorization scale to use running from *not at all* through *barely*, *sort of* and *very much*, to *completely*. If people felt that a category has a clear definition (even if they are uncertain what it is), then we expected them to be disinclined to use a partial rating such as “sort of” or “very much”, and to stick to the two extreme responses—“not at all” or “completely”. In the event our manipulation of context had no effect on this measure either.

In the light of these results, it would appear that instability and disagreement in categorization is not exclusively driven by a lack of specificity in the context. Whether simply classifying, trying to capture common categorization practices, or yet advising a technical committee on the correct way to classify, people rely on the same underlying conceptual representation and this is best described in terms of a typicality gradient.

Typicality is a measure that has had wide use in Psychology, but its relation to semantics is often rather obscure, and subject to misunderstanding. In the next section I therefore discuss in detail just what typicality is measuring, and how it relates to issues of semantics.

1.5 Typicality and Gradedness

In Hampton (2007) I define a position on the relation of typicality to graded categorization. In the Threshold Model I propose that a semantic category is represented by a set of intensional information in the form of a prototype (which may include schematic structure about causal-explanatory links between features). Potential exemplars can be ranked in terms of their similarity to this prototype, as determined by an asymmetric measure of how well the exemplar matches the prototype features. For example, similarity of a tomato to fruit will be greater on this measure than the similarity of fruit to tomato. It is assumed that tomato matches more features of fruit than fruit matches features of tomato because of the greater abstraction of fruit.

The ranking based on similarity then provides the basis for judgments of typicality (assuming other factors are held constant—see below), and also provides the basis for categorization decisions through the application of a criterion or threshold. With the additional assumption that the placement of the threshold is subject to normally distributed error both within and across individuals, a standard psychometric function is obtained relating the probability of a positive categorization to the underlying similarity.

It is perhaps common in some parts of the Cognitive Science community (e.g. Armstrong et al. 1983; Fodor 1998; Osherson and Smith 1981, 1982, 1997) to dismiss typicality effects as purely psychological and hence peripheral to the development of lexical semantic theory. After all, we know that the way in which the mind stores words in the mental lexicon shows all kinds of psychological influences that are orthogonal to issues of lexical meaning. Frequency of a word in the language, for example, has large effects on reading, memory and a range of other cognitive tasks. It has also been found that the degree to which words are associated (like “Fish” and “Chips” in the UK) can be highly predictive of a range of phenomena. Why should not typicality effects be of the same kind?

I will argue that when people judge typicality (for example of an item in a category) they may be judging quite a number of different things. Nonetheless paramount among those different dimensions is similarity to a prototype representing the common intensional properties of the class. Because this measure of similarity also determines the degree to which an item can be said to belong in a category, typicality ratings do an excellent job of predicting the likelihood that an item will be placed in a category as evidenced in the Hampton et al. (2006) experiments, and many other similar studies.

1.6 Does Variation in Typicality Really Undermine the Classical Model?

Any description of the impact of Rosch's prototype theory on the psychology of concepts tends to make much of typicality effects. The classical theory against which Rosch was arguing proposed that a concept could be defined intensionally as a conjunction of individually necessary and jointly sufficient features. A concept such as "bachelor" could be defined as "human, male, adult, and eligible to marry". It has been claimed by supporters of the prototype model that this classical theory gives equal status to all items that meet the definition, so there should be no differences between items in terms of how well they represent the class. Although often repeated, I do not take this to be a fair criticism of the classical model. Typicality can reflect many different underlying structural variables, many of which do not relate to the question of whether a concept term applies to an exemplar. It is important therefore to tease apart the different influences on typicality to get a clear picture of the role that it plays in conceptual structure.

First, the notion of typicality or goodness-of-example is often confounded with other non-semantic dimensions such as familiarity (Malt and Smith 1982). Indeed Armstrong et al. (1983) demonstrated that well-defined categories such as Odd and Even Numbers have clear typicality structure, most probably based on simplicity and familiarity (but see Larochelle et al. 2000, for counter-evidence). (Of course, the demonstration that well-defined categories show typicality effects does nothing to undermine the theory that for other types of categories, lacking an explicit definition, typicality may be a critical factor in determining membership.)

Second, it has been suggested that category membership is determined by a defining core of features, whereas additional "characteristic" features are associated with typicality differences (Osherson and Smith 1981; Rey 1983; Smith et al. 1974). In support of this proposal, Rips (1989) presented a variety of attempts to dissociate measures of similarity, typicality and category membership, in which it is claimed that some item may be more typical of category A than of category B, even though it is a better member of B than of A. His results have however not stood up well under replication (Hampton et al. 2007; Smith and Sloman 1994).

Given that Typicality effects on their own do not provide strong evidence against the classical model, what do we know about their basis?

1.7 Ingredients of Typicality 1: Ideals

It has been shown (e.g. Barsalou 1985) that when people are asked to say how typical an item is as a member of a category, then they are influenced by several different dimensions. First and foremost, Typicality ratings are assumed to be a pure measure of the underlying similarity in meaning, or degree of match of semantic features, between a member and its superordinate category. This dimension of

similarity is clearly a major influence on Typicality ratings. However other factors are also involved. Barsalou (1985) showed that in addition to what he termed Central Tendency (closeness of a concept to the centre of its category), ratings of Typicality were also correlated with frequency of instantiation (a measure of familiarity) and matching of Ideals. An ideal is a feature of a concept that represents extreme rather than average values of a dimension. Thus a winter coat may be considered most typical if it is ideally warm and light, as opposed to being closest to the average winter coat (which will be of average warmth and average weight).

So Typicality per se must be interpreted with this ambiguity in mind. The ambiguity has been made a lot worse by Rosch's (1975) original characterization of typicality as "goodness-of-example", a term also used in Barsalou (1985). Barsalou notes that the word "typicality" was not used in his study because it could bias participants into thinking of frequency of instantiation. As a consequence, he asked for goodness-of-example, with a scale running from "poor example" to "excellent example". Similarly, Burnett et al. (2005) concluded that expert fishermen judged the typicality of types of fish based on ideals, while using "goodness of example" as their measure of typicality. The difficulty here is that asking about "goodness" leads to an evaluative judgment and hence allows ideals to have a greater influence on the judgments.

There have been very few attempts to distinguish between the two senses of typicality. One exception is Kittur et al. (2006) who dissociated the two dimensions using a novel relational concept learned in the laboratory. They found that ratings of goodness of example reflected just ideals, whereas typicality judgments reflected both ideals and central tendencies.

One way to understand the relation of typicality and ideals would be to propose that ideals should be understood as contributing to typicality itself which then determines degrees of membership. In this way, typicality would mediate the influence of ideals on category membership. This proposal needs empirical testing. In an experimental manipulation of ideals, Kim and Murphy (2011) demonstrated that in fact ideal exemplars that best served a category's goals were not necessarily perceived as most typical. For example, a great party might be considered ideal, but was not judged as typical.

1.8 Ingredients of Typicality 2: Frequency and Familiarity

As well as Ideals, Barsalou (1985) also identified Frequency of Instantiation as a component of Typicality. Participants were asked to judge subjectively how often a category member occurred as an instantiation of the category. Allied to this measure is a second measure: Category Dominance. Going back to the early days of associationist psychology, the measure of Category Dominance is the relative frequency with which an item is generated when people are asked to list all the category members that come to mind within a limited time (Battig and Montague 1969). A third measure related to frequency of instantiation is familiarity, in which

participants rate items in terms of how familiar they seem. These different measures tend to correlate together and to form a separate dimension from typicality owing to family resemblance or similarity (Barsalou 1985; Hampton and Gardiner 1983).

Hampton (1997a) was able to show a double dissociation of the effects of Typicality and Category Dominance on response times and errors when people make speeded categorization decisions. In a first experiment, regression methods were used to differentiate the effects of typicality, familiarity and category dominance on the average time to categorize items, and the likelihood of making a positive response. Participants were given a category (e.g. Fruit) and then a list of words one at a time. They had to make a speeded decision for each word whether it belonged in the category or not. Mean reaction time was predicted in a multiple regression equation using norms for semantic categories collected by Hampton and Gardiner (1983). (Hampton and Gardiner 1983, used instructions for typicality that explicitly differentiated it from frequency of instantiation.) Likelihood of a Yes response was also predicted. The results showed that typicality and category dominance each made independent and significant contributions to predicting decision time. Although the two measures were correlated with each other, the speed in making a decision was driven both by the availability of the item in memory (as measured by category dominance) and by the similarity of the item to the category (as measured by typicality). When it came to response probability, only typicality predicted the likelihood of a Yes or No response.

The second experiment introduced two manipulations. First, manipulating the difficulty of the task by including closely related false items (e.g. a bat is a bird) slowed down atypical items relative to typical items, but had no effect on items as a function of their category dominance. A second manipulation in which half the items were seen in a different context before having to be categorized showed that the category dominance effect but not the typicality effect was eliminated by earlier exposure of the items. Taken together the results all suggest that while high dominance items are more readily available in memory, the actual decision of whether something is in a category is just affected by its typicality and not by associative strength. There is an interesting parallel here with the heuristics of Availability and Representativeness proposed by Tversky and Kahneman to explain people's judgments of probability (Kahneman et al. 1982).

1.9 Typicality and Membership

Having described the multiple influences on typicality, including the ambiguity of what it is to be a "good" example, and the confounding with familiarity and category dominance, what is the evidence that there is nonetheless a purer notion of typicality that should be taken seriously as a component of meaning? The results of the study by Hampton et al. (2006) described above provide one such piece of evidence. Here we had eight categories in which there was uncertainty or vagueness in the classification of borderline items. Moreover, the basis of the disagreement

and inconsistency did not come down to mere ignorance. There is no fact of the matter or correct answer to these borderline categorization cases. In that respect they can be termed an example of Vagueness, with similar properties to the traditional cases of vagueness seen in adjectives like red, bald or tall. Being bald is a matter of degree, so that it is meaningless to ask exactly how many hairs need to be lost before someone is correctly termed bald. In the same way, many noun categories have membership which is a matter of degree (sociology may be considered a better science than palm-reading) but there is no hard and fast way to determine who is right and who is wrong in the event of a borderline dispute. We come back again to the dispute about bridge being a sport. With due respect to the UK Supreme Court, there is no higher authority to which one can turn to decide the question in an objective fashion (as there might be in the case of a biological or technical term). As in matters of taste, it appears that each may be entitled to his/her own opinion about such cases (Wright 1995).

It is in the context of this vagueness in noun meanings that the notion of Typicality can be helpful. Disputes about borderline cases often end up with party A arguing “X is a sport because it has features D, E and F”, while party B argues “X is not a sport because it lacks features P, Q and R”. But this is exactly what the “pure” notion of Typicality captures—the fact that the more features of a concept an item possesses the more justified one is in placing it in the category. There is a continuity between one category member being more typical of a category because it has more matching features (as in the case of a robin being a more typical bird than an ostrich, even though both are clearly birds) and one item being more likely to be classed as a member of a category than another for the same reason (as in ten-pin bowling being considered more of a sport than billiards.)

This is a critical point for the debate between formal semantics and psychology. If typicality is a purely psychological phenomenon that does not affect truth values (as in the robin versus ostrich case) then it is safe for semantics to ignore it and instead to focus on category membership (a position taken among others by Osherson and Smith 1997). However when other conceptual categories are considered, it turns out that variation in typicality (as determined by similarity to a prototype or by the degree to which something possesses the prototypical features) does affect truth values. In the Hampton et al. (2006) experiments we showed again and again that rated typicality was the best predictor of people’s judgments of truth for sentences such as “seaweed is a vegetable”, “a tomato is a fruit”, “a squid is a fish” or “a piano is a kind of furniture”.

In Hampton (2007) I argued therefore for a single underlying dimension that in the first place determines how typical some item is of its class (in the family resemblance sense of typical, rather than anything involving ideals or familiarity), and in the second place determines how much of a member of the class it is. This underlying dimension relates to the degree to which the conceptual representation of the item brought to mind in the given task context matches that of the superordinate category that it is being compared to.

To develop this model, the notion of Typicality has to be extended or adapted a little further. Linguistically speaking, to say that X is a typical (or atypical) Y carries the presupposition that X is indeed a Y. One would probably not say that a bat is an atypical bird. It may resemble a bird but that doesn't make it an atypical one. However in the original introduction of typicality ratings as "goodness-of-example" Rosch (1975) also chose to ignore this refinement and asked participants to judge typicality across a full range of items running from typical exemplars through borderline cases to clear non-examples. Adapting to the pragmatics of the task, her participants duly obliged, and so Typicality also has an extended meaning corresponding to something like "typicality if it is a member and closeness to the category if it is not".

In a number of papers, (Hampton 1979; Hampton and Gardiner 1983; Hampton 1988b) I developed a graded membership scale based on this more explicit notion which was then used in the studies of concept conjunction negation and disjunction described above. As discussed above, the key result of those studies was that membership in a conjunctively defined category showed a continuous gradation (in terms of the probability of a positive response) that was highly predictable in a regression from degree of membership in the two constituents. Furthermore, just as typicality in a conjunction is known to sometimes surpass typicality in a conjunct (the guppy as a pet fish is the example proposed by Osherson and Smith 1981), so **membership** in a conjunction can surpass membership in a conjunct.

This phenomenon of overextension has been demonstrated most recently in a study of activity verbs by Martin Jönsson. Jönsson (2015) showed people videos of an actor simultaneously performing two actions, such as Smoking and Walking. The action would be a typical example of the first action (e.g. smoking) but a very atypical example of the second (e.g. walking). The task asked for a Yes/No answer to simple questions such as "Is this man smoking?", "Is this man walking?" or "Is this man smoking and walking?". Jönsson found with this particular example that 100% answered yes to the first question, and only 39% answered yes to the second. However 70% answered yes to the conjunctive question. It appears that likelihood of agreeing to classification in a conjunction may involve an average of perceived degrees of membership in each conjunct, rather than the likelihood of believing that the item is in the first category and also believing that the item is in the second.

A crucial test of the involvement of Typicality in categorization judgments is to show that variation of typicality among cases which are **clearly members of a category** can nonetheless affect categorization in a conjunction. Suppose that an item is clearly a member of one category, but is on the borderline for another. For example, suppose that everyone agrees that Jack is bald (about half his head is hairless) but only 50% agree that he is tall. Then the composite prototype account would predict that increasing Jack's degree of baldness yet further would compensate for his lack of tallness and make it **more likely** that people would accept that he is both bald and tall.

This prediction of compensation between typicality in one category and membership in a conjunction was tested in Hampton (1996). People made a set of three judgments about cartoon faces representing a range of age from child to adult and a

range of emotion from happy to sad. Critical test items were clearly either children or adults, but differed in typicality in those categories. At the same time these test faces were borderline in terms of emotion. The study showed, for example, that variability in the typicality of a given adult face influenced the likelihood of categorization in a conjunction such as “happy adult”. Thus typicality of a face as an adult face could compensate for borderline membership as a happy face and affect the categorization in the conjunction.

1.10 Differentiating Vagueness from Ignorance

A recent paper (Hampton et al. 2012) shows that the ontological uncertainty about whether (say) bridge is a sport can be differentiated from other kinds of uncertainty based on ignorance. In a set of studies we looked at the problem of higher order vagueness. The basic set up was as follows. Two groups of people had to judge whether a list of items belonged in a category, and as previously, the list was designed to contain many borderline or disputable items. The procedure required them to return after two weeks and to perform the task again (as in McCloskey and Glucksberg 1978). We measured the likelihood that they would give the same response on each occasion—a measure we labelled as Consistency. In the first group who responded simply Yes or No, people would typically maintain the same response for about 80% of items. People in the second group, rather than saying simply Yes or No, were given the chance to create a third, middle, response category. They were instructed to first decide if the item was 100% certain to be in the category, or 100% certain to be NOT in the category, and respond accordingly. Any item for which they were not 100% sure, they were told to put in a middle response category of “Not 100% sure”. As for the first group, we measured consistency of responding.

Our original intuition was that this second condition would lead to greater consistency. As people could “cherry pick” the easy items and leave the others aside, and as they would be given credit for being consistent if they put an item in the “Not 100% sure” category both times, we felt that they would change their minds much less often. In fact, the results were quite clear in showing that the likelihood of changing your mind about whether something is 100% certain to be in a category is no more nor less than the likelihood of being inconsistent in judging if it is in the category or not. Higher order vagueness (determining the boundaries of the vague region where things are unsure) turned out to be just the same as lower order vagueness (the indecision within the vague region).

This effect can be used to argue that vagueness about category membership is not equivalent to uncertainty owing to ignorance. In different versions of the task, we tried the same procedure with general knowledge statements. Instead of a statement such as “rhubarb is a fruit” we had statements like “The Uruguayan flag has red in it”. Now, the second group who were allowed to say when they were unsure were significantly more consistent in their responses. After several other

studies, including both those in the paper and other unpublished studies done since, we can conclude that when there is an objective truth to a statement (as in general knowledge, or as in the correct meaning of a word, or as in memory for a video) then people can reliably identify the statements that they “know they don’t know”. On the other hand, when the truth of a statement has a more subjective basis, as in categorization but also as in personal judgments about one’s aspirations, moral beliefs or early childhood memories, then one cannot do so. In these cases, asking people to only say “yes” when they are definitely sure simply moves the decision criterion to a higher level but does nothing to reduce the inherent unreliability of the decision.

A recent unpublished study, conducted with Shauna-Kaye Williams demonstrates this effect with a single dimensional example of vagueness. We had two sets of faces which varied according to their emotional expression. The first were a set of morphs between a neutral expression and a happy expression. The second were morphed between a neutral expression and a surprised expression. In a first session, participants went through each set of faces twice, once deciding if the faces were happy or not, or surprised or not (depending on the set), and once deciding if they were *clearly* happy or not, or *clearly* surprised or not. They returned for a second session a week later, and repeated the task. In this case, we were interested in whether asking people only to respond positively if the faces were “clearly” showing the emotion would lead to a sharper boundary and more consistent responding over time. In fact, neither of these occurred. Figure 1 shows the results where logistic regression functions were fit to each individual’s data and then a plot made based on the average slope and threshold. In both sets of faces, the requirement to select faces that clearly showed the expression simply moved the threshold to the right, while leaving the sharpness of the boundary unaffected. (In fact, the surprised faces had a slightly lower slope—a vaguer boundary—for the clearly judgments).

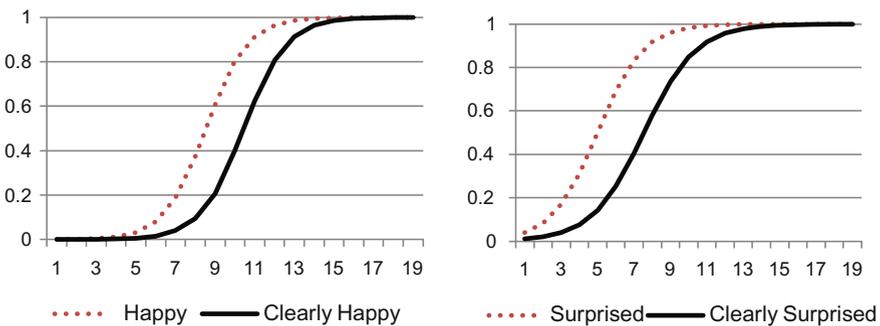


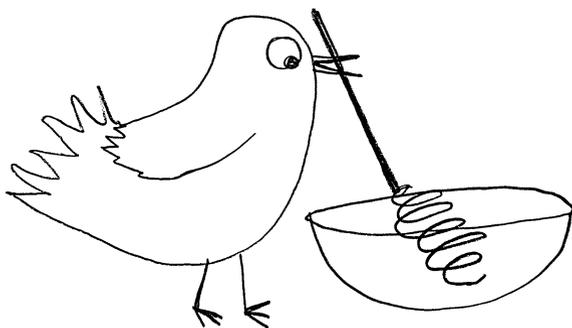
Fig. 1 Logistic functions fit to the likelihood of a positive categorization as a function of the morphed scale of emotional intensity for happy or clearly happy faces (*left*) and surprised or clearly surprised faces (*right*)

1.11 *Concept Intensions as Fundamental*

I have hoped to show that there are a number of phenomena that argue for a fundamental role for intensional representations of concepts as the raw components of thoughts. My research into conjunctions of prototype concepts shows that people are not particularly concerned to respect the logic that requires that to be in a conjunction means to be in each of the conjuncts. Likewise there is good evidence that neither disjunction nor negation fares any better in terms of maintaining logical norms (Hampton 1988a, 1997b). To account for these deviations from logic, it is most fruitful to look at how intensions might be combined in different ways to form conjunctive, negated or disjunctive concepts (Hampton et al. 2012). The Composite Prototype Model (CPM, Hampton 1987, 1988b) details how we might account for the results through a process of integrating two prototype concepts into a single composite to represent the complex concept. As Pelletier (2017) describes it, this is compositionality as applied to prototypes. We know that prototypes for conjunctions (e.g. pet birds) may look quite different from those of their conjuncts (pets and birds considered separately). As described above, the model suggests an initial process of attempting to combine the two concepts by taking the disjunction of their features. This is followed by the identification of points of incompatibility and either the deletion of certain features (pet birds do not migrate in winter) and/or the addition of new emergent features not seen in either concept (pet birds live in cages though neither birds nor pets do normally).

In order to explore the process of conceptual combination in a “pure” state, without the influence of prior familiarity, Hampton (1997c) described a study in which people were required to describe conjunctions of categories which do not currently exist, such as a computer that was also a kind of teacup, or a vehicle that was also a kind of fish. Why set people such a task? I was primarily interested in demonstrating that concept intensions are flexible and adaptable. Just as when forming the concept of Pet Fish one must abandon some salient features of each concept (we don’t eat pet fish with French fries, and nor do they cuddle on our laps), so when truly incompatible conjunctions need to be imagined the process can be taken to extremes. Figure 2 shows an example of one of the solutions offered by

Fig. 2 A bird that is a kitchen utensil. Redrawn from original drawing of an anonymous participant



the more creative participants, in response to the requirement to find a bird that was also a kitchen utensil. In this case a woodpecker has been trained to whisk eggs using its powerful head movements.

Qualitative analysis of the results of this first pilot study showed some very interesting principles were involved in combining concepts. First, there was evidence for instantiation of superordinate categories to basic level categories—birds became woodpeckers, fruit bananas, and furniture couches. Second, people would align properties and functions between the two concepts. The example given here shows how the need to find a function for a kitchen utensil is met by finding a behavior of a particular bird that can serve the function of a particular utensil. Third, there was evidence for simulation processes, in line with Barsalou's suggestions about the role of concepts in achieving goals (Barsalou 1991). Commentary provided by the creator of the woodpecker whisk pointed out that it would not need electrical power (good for camping trips) but would on the other hand be unhygienic. The use of simulation takes the specification of the combined concept and develops it in totally non-compositional directions, as the new concept is adapted to real world knowledge. Finally, many solutions identified conflicting properties in the newly combined concept, and offered new emergent features to resolve them. For example, when a participant solved "A fruit that is also a kind of furniture" by proposing a banana couch, they went on to specify that the banana had been modified to grow very large and to ripen very slowly.

Extensional accounts of meaning focus on the sets of exemplars in the world. As such they can have nothing to say about concepts which have no members. (The problem of empty and fictional names is well known, Braun 2005). Thus when we think about counter-factual or hypothetical objects, the extension has to be taken to refer not to the actual world but to a virtual or "possible" world. Computer teacups do not currently exist, but it is possible to imagine an alternative world in which they do, and we can speculate about their properties. In fact this process of conceptual combination may be a key factor in innovation and creativity.

In a recent paper (Gibbert et al. 2012) my colleagues and I investigated creativity in the forming of concepts of hybrid products—artifacts that serve more than one traditional function. In direct agreement with the processes postulated by the Composite Prototype Model we showed how first attempts to imagine (say) a pillow that was also a telephone would simply aggregate the features of the two concepts into a composite. However when the two concepts being combined were sufficiently dissimilar, then a second attempt at combining the concepts would generate integrative solutions in which features of one concept would be aligned with those of the other to provide emergent functionality. The telephone could be programmed to provide gentle sounds to help people get to sleep, or the pillow could gently move to alert the user that a phone call was arriving. A less kind proposal was that the pillow would allow one to nap while listening to one's mother on the phone. When people were able to generate integrative solutions, these were consistently judged to be more likely to succeed as marketable products.

1.12 Impossible Objects and Hierarchical Levels

To conclude this chapter I will describe a study in this project, conducted with Diane Lewis and Zachary Estes in which we looked more specifically at the issue of instantiation. We designed a study in which half the concepts were basic level categories (Rosch et al. 1976) corresponding to the most common names for objects, like CAR and LEMON and the other half were superordinate categories like VEHICLE or FRUIT, where there is no single image than can be formed of the category as a whole.¹ We had two ideas in mind about how this variable would affect the likely success of forming incompatible conjunctions. On the one hand, superordinate categories place fewer constraints on the solution. For an object to be a car requires some minimum description in terms of shape, size and material which need not be true for something to be a vehicle. We therefore predicted that in general our participants would find it easier to combine disjoint sets at the superordinate level, leading to Superordinate-Superordinate combinations being the easiest to form, Basic-Basic combinations to be hardest, and mixed combinations in between. As a rider to this hypothesis, we also considered it likely that successful use of Superordinate concepts would be found greatest when they were actually instantiated in the solution as familiar basic level concepts.

Set against this prediction was the possibility that because basic objects are very familiar and easily imagined, there would be some processing advantage to having the task set at this concrete level. Indeed there is evidence (Smits et al. 2002) that people actually make decisions about superordinate categories by retrieving prominent basic level exemplars. In particular, we felt that a case could be made for arguing that the mixed conditions Superordinate-Basic and Basic-Superordinate would be the easiest. In these conditions one of the concepts would be anchored to a familiar and concrete basic level object, and the task would then be to modify or transform this easily imagined object in order to meet the criteria of the other superordinate category in some way.

The design was a 2×2 within subjects design manipulating whether the first and second nouns were basic or superordinate concepts. Thirty-two students and other young adults (18 female) aged 17–32, completed booklets. Eight sets of items were constructed. Each set was a quadruple of a pair of basic level concepts to be combined, together with their superordinates. Table 2 lists the materials used. A combination was created by taking one of the nouns from the left two columns and combining it with one of the nouns from the same row in the right two columns. There were thus four possible combinations for each quadruple, such as Banana-Bus, Banana-Vehicle, Fruit-Bus and Fruit-Vehicle.

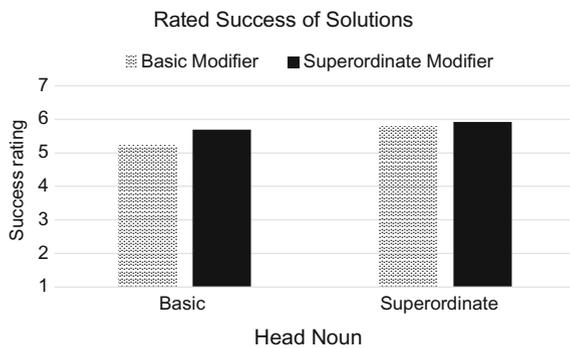
Each participant was given two pairs in each of the four conditions, and the materials were rotated through the four conditions across four groups of participants

¹Some of our superordinate concepts, like Bird and Fish were probably at the basic level in terms of Rosch et al.'s (1976) criteria. However the manipulation of level was still valid with "Woodpecker" providing greater constraints on a combined concept than "Bird".

Table 2 Materials used in the experiment for forming “impossible” combinations of concepts

First noun		Second noun	
<i>Basic</i>	<i>Superordinate</i>	<i>Basic</i>	<i>Superordinate</i>
Banana	Fruit	Bus	Vehicle
Woodpecker	Bird	Jumper	Clothing
Shark	Fish	Daffodil	Flower
Table	Furniture	Mosquito	Insect
Bomb	Weapon	Carrot	Vegetable
Snake	Reptile	Bungalow	Building
Horse	Mammal	Spanner	Tool
Hamster	Pet	Guitar	Musical instrument

Fig. 3 Success of solutions as a function of level of the head and modifier nouns



so that each quadruple contributed equally to all conditions. The first nouns were always the Head nouns, and the second the Modifiers in a phrase such as “A banana that is also a bus”.

1.13 Results

Responses were rated by two independent judges on scales of success (1–10) and symmetry (rescaled for analysis as 1 = bias towards head, 0 = no bias, –1 = bias towards modifier). Reliabilities were estimated by the Spearman Brown method as 0.70 for success ratings and 0.75 for symmetry. Raters disagreed by more than 4 scale points for the success rating on only 5% of occasions, and gave opposing symmetry judgments also on only 5% of occasions.

Success. Figure 3 shows the mean rated success of solutions as a function of the level of the head and modifier concepts. Problems involving superordinate terms were more successfully solved than those involving basic level terms, both for the head noun (5.8 vs. 5.4) and for the modifier noun (5.8 vs. 5.5). ANOVA confirmed independent significant main effects of level for the head noun ($F(1, 31) = 6.9$,

$p < 0.05$, and for the modifier noun ($F(1, 31) = 4.55, p < 0.05$), with no significant interaction ($F < 1$).

Symmetry. Analysis of the Symmetry judgements showed that solutions tended to be more similar to the head noun than the modifier (mean bias = 0.14, SD = 0.54, $t(127) = 2.84, p < 0.005$), but there were no effects of level. This result suggests that the head noun may be taken as a starting point and modified in the direction of the modifier, rather than vice versa.

Instantiation. The significant effects of level on success were in keeping with the prediction that the greater flexibility allowed to the participant from the use of a superordinate would allow better solutions to be found. To test this notion further we examined the interaction between success and the amount of instantiation used in a solution.

Solutions for the conditions involving superordinates (Basic-Superordinate, Superordinate-Basic and Superordinate-Superordinate) were divided on the basis of whether either the head or modifier superordinate was clearly instantiated as a particular basic level term. Some of the 16 superordinate categories were almost always instantiated (notably furniture, fruit, mammal, pet, vehicle) whereas others were almost never instantiated (flower, fish). The likelihood of a superordinate concept being instantiated was greater (72%) when the other concept was also superordinate, than when the other was a basic level term (59%), so it appears that finding a solution is easier if **at least one** of the terms is at the basic level or has been instantiated at the basic level.

Average success of solutions was compared for cases where the superordinate(s) were instantiated and cases where they were not. For this purpose instantiation was treated as a Post Hoc Factor. For the conditions with one Basic and one Superordinate concept, instantiation of the superordinate had a relatively small effect (mean success = 6.0 instantiated, 5.4 un-instantiated, $t(126) = 1.97, p = 0.051$).

However for the Superordinate-Superordinate condition, instantiation had a sizeable effect on success. Where neither was instantiated (13 out of 64 cases), mean success was only 4.3, whereas when either one or both were instantiated it rose to 6.3. Because of small cell sizes, a one-way ANOVA was conducted comparing three levels of instantiation—both nouns instantiated, just one instantiated, or neither noun instantiated. There was a strong effect of instantiation ($F(2, 61) = 6.1, p < 0.005$). Figure 4 illustrates some of the more successful solutions offered.

What has been shown by these explorations of the creative potential in our conceptual system? Clearly the primary function of words and sentences is to enable us to communicate and coordinate our thoughts about the world. Lexical items have a dictionary meaning that provides a firm basis for learning a language and using it effectively across a range of social contexts. At the same time, the concepts that constitute the basis of those meanings are capable of showing a flexibility that is fundamental to the process of invention and conceptual change. To understand how lexical composition occurs it helps to understand more about how conceptual contents can be combined. That process would appear to require access to the full repertoire of human cognitive capacities, well beyond the limits of a set

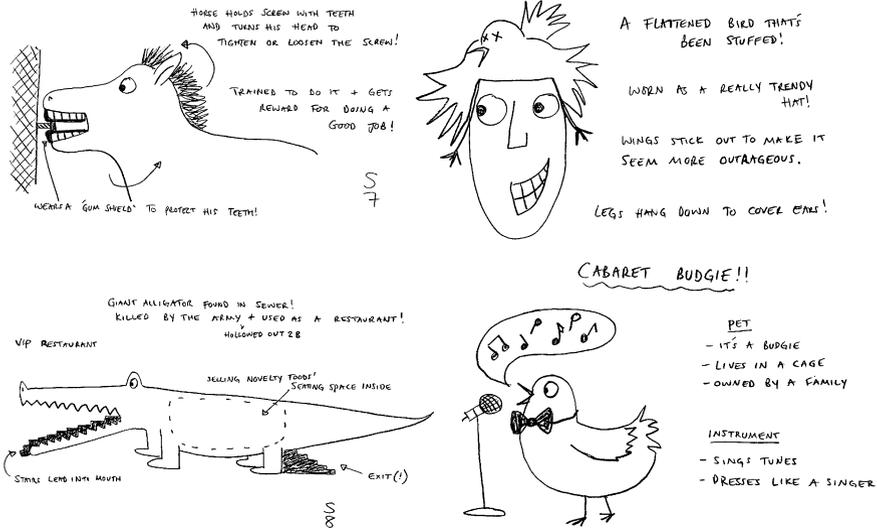


Fig. 4 Examples of the successful solutions generated by the more creative participants. Clockwise from *top left* a horse which is a tool, a bird which is clothing, a pet which is a musical instrument, and a reptile which is a building. Instantiation, alignment and emergent features are evident in these solutions. Original drawings of anonymous participants have been redrawn

of compositional rules applied to a finite set of fixed concepts. In particular it may require active simulation of concepts in an imagined world (Barsalou, 2017) and access to knowledge of the world (Murphy and Medin 1985).

2 Conclusions

In this chapter I have tried to provide some hard evidence that lexical meanings are often concepts that are constituted as prototypes in the mind of the language user. The prototype consists of a set of correlated features that represent what people know and expect about the most typical or representative example of the kind, and the amount of variability that can be expected around that. Because these concepts lack hard definitions, it is common to find borderline disputes about meaning. Furthermore when modelling the way in which people interpret apparently conjunctive phrases such as “an A which is a B” it is necessary to take account of these prototype intensions in order to explain the patterns of overextension and compensation that occur, together with other effects such as non-commutativity and category dominance.

I hope also to have provided an explanation of why typicality should not be taken as a unitary measure, since as a task it can invite a range of different pragmatic interpretations, including ideals and familiarity as well as the intended one of representativeness.

Let me conclude with some comments about the relation of this work to the work of formal semantics as described in the chapter in this volume by Pelletier. It is perfectly true that the work I have described offers no account of the difference between individuals and kinds, no account of how the scope of quantifiers is determined from syntax and no account of indefinitely many other linguistic and semantic phenomena. That was never its aim. The focus on intersective noun combinations was primarily to demonstrate that you need intensions to explain people's intuitions of applicability in these cases. There is a large psychological literature on other forms of conceptual combination involving noun-noun compounds (Wisniewski 1997; Gagne and Shoben 1997; Estes and Glucksberg 2000) showing that intersective interpretations are relatively rare compared to a number of other commonly used thematic relations such as "MADE OF" (e.g. CHOCOLATE EGG), "LOCATED IN" (e.g. CITY BUS) or "USED FOR" (e.g. CEMENT TRUCK). The question is therefore whether the two distinct approaches to semantics can find a way to mesh, or whether there are fundamental incompatibilities between them. A problem of finding the right conceptual combination.

References

- Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, 13, 263–308.
- Barsalou, L. W. (1991). Deriving categories to achieve goals. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 27, 1–64.
- Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 629–654.
- Barsalou, L.W. (2017). Cognitively plausible theories of concept composition. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 9–30). Berlin: Springer.
- Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, 80(3, part 2).
- Bowerman, M., & Choi, S. (2003a). Space under construction: Language specific spatial categorization in first language acquisition. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition* (pp. 387–428). Cambridge: MIT Press.
- Bowerman, M., & Choi, S. (2003b). Space under construction: Language specific spatial categorization in first language acquisition. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought*. Cambridge MA: MIT Press.
- Braun, D. (2005). Empty names, fictional names, mythical names. *Noûs*, 39, 596–631.
- Burnett, R., Medin, D., Ross, N., & Blok, S. (2005). Ideal is typical. *Canadian Journal of Psychology*, 59, 3–10.

- Davies, Caroline (2015). High court ruling paves way for courts to decide whether bridge is a sport. *The Guardian*, 27/04/2015.
- Estes, Z., & Glucksberg, S. (2000). Interactive property attribution in concept combination. *Memory and Cognition*, 28, 28–34.
- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford: Oxford University Press.
- Fodor, J. A., & Lepore, E. (2002). *The compositionality papers*. Oxford University Press.
- Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of Modifier-Noun combinations. *Journal of Experimental Psychology: Learning Memory, and Cognition*, 23, 71–87.
- Gibbert, M., Hampton, J. A., Estes, Z., & Mazursky, D. (2012). The curious case of the refrigerator-TV: Similarity and hybridization. *Cognitive Science*, 36, 992–1018.
- Goodenough, W. H. (1965). Yankee kinship terminology: A problem in componential analysis. *American Anthropologist, New Series*, 67(5), 259–287.
- Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 18, 441–461.
- Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, 15, 55–71.
- Hampton, J. A. (1988a). Disjunction of natural concepts. *Memory & Cognition*, 16, 579–591.
- Hampton, J. A. (1988b). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 14, 12–32.
- Hampton, J. A. (1996). Conjunctions of visually-based categories: Overextension and compensation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 22, 378–396.
- Hampton, J. A. (1997a). Associative and similarity-based processes in categorization decisions. *Memory & Cognition*, 25, 625–640.
- Hampton, J. A. (1997b). Conceptual combination: Conjunction and negation of natural concepts. *Memory & Cognition*, 25, 888–909.
- Hampton, J. A. (1997c). Emergent attributes in combined concepts. In T. B. Ward, S. M. Smith, & J. Viad (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 83–110). Washington DC: American Psychological Association Press.
- Hampton, J. A. (1998). Similarity-based categorization and fuzziness of natural categories. *Cognition*, 65, 137–165.
- Hampton, J. A. (2007). Typicality, graded membership and vagueness. *Cognitive Science*, 31, 355–383.
- Hampton, J. A., Aina, B., Andersson, J. M., Mirza, H., & Parmar, S. (2012). The Rumsfeld effect: The unknown unknown. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 38, 340–355.
- Hampton, J. A., Dubois, D., & Yeh, W. (2006). The effects of pragmatic context on classification in natural categories. *Memory & Cognition*, 34, 1431–1443.
- Hampton, J. A., Estes, Z., & Simmons, S. (2007). Metamorphosis: Essence, appearance and behavior in the categorization of natural kinds. *Memory & Cognition*, 35, 1785–1800.
- Hampton, J. A., & Gardiner, M. M. (1983). Measures of internal category structure: A correlational analysis of normative data. *British Journal of Psychology*, 74, 491–516.
- Johnson, K. (2004). On the systematicity of language and thought. *The Journal of Philosophy*, 101 (3), 111–139.
- Jönsson, M. L. (2015). Overextension in verb conjunctions. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 41(6), 1917–1922.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kim, S., & Murphy, G. L. (2011). Ideals and category typicality. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 37(5), 1092–1112.
- Kittur, A., Holyoak, K. J., & Hummel, J. E. (2006). Ideals aren't always typical: Dissociating goodness-of-exemplar from typicality judgments. In R. Sun & N. Miyake (Eds.), *Proceedings*

- of the Twenty-eighth Conference of the Cognitive Science Society (pp. 429–434). Mahwah, NJ: Erlbaum.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.
- Landau, B. (1982). Will the real grandmother please stand up? The psychological reality of dual meaning representations. *Journal of Psycholinguistic Research*, 11(1), 47–62.
- Larochelle, S., Richard, S., & Soulières, I. (2000). What some effects might not be: The time to verify membership in “well-defined” categories. *Quarterly Journal of Experimental Psychology, A*, 53, 929–961.
- Malt, B. C., & Smith, E. E. (1982). The role of familiarity in determining typicality. *Memory & Cognition*, 10, 69–75.
- McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, 6(4), 462–472.
- Murphy, G. L., & Andrew, J. M. (1993). The conceptual basis of antonymy and synonymy in adjectives. *Journal of Memory and Language*, 32, 301–319.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289–316.
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9(1), 35–58.
- Osherson, D. N., & Smith, E. E. (1982). Gradedness and conceptual combination. *Cognition*, 12(3), 299–318.
- Osherson, D. N., & Smith, E. E. (1997). On typicality and vagueness. *Cognition*, 64, 189–206.
- Pelletier, F. (2017). Compositionality and concepts—A perspective from formal semantics and philosophy of language. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 31–94). Berlin: Springer.
- Pullum, G. K., & Scholz, B. C. (2007). Systematicity and natural language syntax. *Croatian Journal of Philosophy*, VII(21).
- Rey, G. (1983). Concepts and stereotypes. *Cognition*, 15, 237–262.
- Rice, S. A. (1992). Polysemy and lexical representation: The case of three English prepositions. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 89–94). Hillsdale, NJ: Erlbaum.
- Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 21–59). Cambridge: Cambridge University Press.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–232.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.
- Smith, E. E., & Sloman, S. A. (1994). Similarity-versus rule-based categorization. *Memory and Cognition*, 22, 377–386.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A feature model for semantic decisions. *Psychological Review*, 81, 214–241.
- Smits, T., Storms, G., Rosseel, Y., & De Boeck, P. (2002). Fruits and vegetables categorized: An application of the generalized context model. *Psychonomic Bulletin & Review*, 9, 836–844.
- Storms, G., Ruts, W., & Vandenbroucke, A. (1998). Dominance, overextensions, and the conjunction effect in different syntactic phrasings of concept conjunctions. *European Journal of Cognitive Psychology*, 10, 337–372.
- Werning, M., Hintzen, W., & Machery, E. (2012). *The Oxford handbook of compositionality*. Oxford University Press.
- Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, 4, 167–183.
- Wright, C. (1995). The epistemic conception of vagueness. *The Southern Journal of Philosophy*, 33, 133–160.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Typicality Knowledge and the Interpretation of Adjectives

Choonkyu Lee

Abstract In this paper, we discuss our experimental results involving color preference and yes/no categorization judgments that provide insight into the interpretation of color adjectives. We selected a set of object categories that show a consistent color typicality bias, and presented them with varying degrees of color manipulation in our experiments. In Experiment 1, Dutch speakers performed a forced-choice picture-phrase matching task. Between a photograph of an object in its typical color (e.g., a light green tomato) and another photograph in a focal color (a darker green tomato), participants showed a significantly higher proportion of preference for the typical, nonfocal color for categories with a color bias (e.g., tomato; 52%) than for categories without a bias (e.g., box; 36%). In Experiment 2, we conducted a categorization task in which participants judged whether an image was an example of the target adjective-noun combination or not, in yes-no format, for 14 adjective-noun combinations including color and other adjectives, such as pattern and material adjectives. When presented with nonfocal images, participants were much more likely to give ‘Yes’ responses for categories with a typicality bias (55%) than for those without a bias (27%), demonstrating an effect of world knowledge in yes-no categorization judgments as well.

1 Introduction

The domain of colors was discussed frequently in the early literature on concept composition (e.g., Smith et al. 1988; Medin and Shoben 1988), and also played a key role in the development of prototype theory (Heider 1972; Berlin and Kay 1969). It continues to be central in the literature on top-down influences on categorization (e.g., Mitterer et al. 2009). In the semantics literature, however, color adjectives have usually been considered intersective adjectives whose composition with nouns simply involves set intersection and whose truth-conditions are inde-

C. Lee (✉)

Department of Linguistics, University of Potsdam, Potsdam, Germany
e-mail: chlee@uni-potsdam.de

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_5

123

pendent of the context they appear in (e.g., Keenan and Faltz 1985; Chierchia and McConnell-Ginet 1996; Drašković et al. 2013).

Previous discussion of adjective-noun combinations involving color adjectives revealed some general issues in meaning composition. Smith et al. (1988) used examples such as *red apple* and *brown apple* to illustrate how their model of meaning composition, ‘selective modification,’ worked. In this model, a noun representation has multiple attributes, including its color (when it is relevant), and adjectival modification of the noun with a color term selectively influences the color attribute. A mental representation of apple thus includes relevant attributes such as its color, shape, texture, and freshness, each with default values set according to our world knowledge. In their example, Smith et al. (1988) assigned the values 25 for red, 5 for green, and 0 for brown in the color attribute of *apple* to model our expectation about different possibilities of the color of an apple. In this model, modifying the noun *apple* with a color adjective *red* or *brown* simply re-assigns the color-attribute values so that the corresponding color gets all the values (30 red, 0 green, 0 brown for *red apple*; and 0 red, 0 green, and 30 brown for *brown apple*), while also increasing the diagnosticity of the color attribute in similarity judgments. When Smith et al. (1988) tested their model’s predictions against participants’ typicality ratings of various items for concepts such as *red vegetable* and *red fruit*, they obtained correlations between 0.70 and 0.97, which were in a similar range to other adjective-noun combinations such as *round vegetable* and *round fruit*.

Medin and Shoben (1988), however, pointed out the limitations of Smith et al.’s (1988) selective modification model with an emphasis on correlated attributes in the world. In the *brown apple* example above, the brown color usually indicates that the apple is not fresh, and neglecting such world knowledge would be an oversimplification for a model of conceptual representation. Participants in Medin and Shoben’s (1988) typicality rating study demonstrated their awareness of correlated attributes, judging white clouds to be more harmless than gray clouds, and black-and-white TVs to be more likely to be small than color TVs. Moreover, their similarity judgments demonstrated that the internal structure of color space changes according to the noun context, with the color pair *white-gray* judged to be more similar than the pair *gray-black* in the context of *hair*, but the other way around in the context of *cloud* or *bear*.

1.1 *Overextension and World Knowledge*

Aside from the impact of world knowledge, Hampton (1996) showed non-Boolean processes in concept composition, using letters in different colors, but there is evidence in Smith and Osherson’s (1984) early data that suggests world knowledge also constrains these processes in concept composition. Hampton (1996) conducted categorization experiments on ambiguous-colored letter shapes, whose values on

the dimensions of color and shape similarity to the letters, A or H, varied along 11-point scales. In one experiment, participants gave a forced-choice response between the colors, blue and green, to judge the color of a letter shape, and later gave a yes-no categorization response for the same visual stimuli with regard to adjective-noun descriptions such as *Blue H*. Hampton (1996) observed that in about 10–20% of the trials, participants accepted descriptions such as *Blue H* after choosing the color *Green* over *Blue* for the same image, when the stimulus color was slightly closer to green but the stimulus shape was a very good H. Another experiment using the colors, orange and red, and yes-no questions such as *Is this Orange?* instead of a forced choice between two colors, showed a similar rate of overextension. The main issue in this study was not world knowledge, as there is presumably no clear color preference between blue versus green As and Hs, or between orange versus red As and Hs in our world knowledge. Hampton's (1996) results showed that categorization judgments that are negative for a single dimension (simply color) can be positive when the dimension is combined with another in which the stimulus has high 'goodness-of-membership,' revealing non-Boolean composition.

Smith and Osherson's (1984) typicality rating results (Experiment 2, p. 349) showed a similar 'relaxation' of standards in the interpretation of adjectival modifiers: In response to pictures that were intermediate between two adjective concepts (e.g., an ambiguous red-brown apple drawing), participants' typicality ratings of the drawing as an example of the adjective-noun description (*red apple* and *brown apple*) were significantly higher than those with just the adjective category (*red* and *brown*). Because Smith and Osherson used concepts for which we have stronger knowledge-based biases than colored letter shapes in Hampton's (1996) study, we can observe the impact of world knowledge on judgments involving complex concepts. When the color in the verbal description was an atypical color for the depicted category in the drawing and thus negatively diagnostic (e.g., *brown apple* and *red canary*), there was always reliable overextension in typicality ratings, namely, significantly higher ratings for the adjective-noun descriptions (*brown apple*, *red canary*) compared to just the color adjective (*brown*, *red*), regardless of the actual degree of match between the image and the description. In contrast, when the color of a description was a typical one for the noun category and thus positively diagnostic (*red apple* and *yellow canary*), there was not consistent overextension of a color judgment from simple descriptions (*red*, *yellow*) to complex ones (*red apple*, *yellow canary*). In fact, when the drawing was a poor match (e.g., a drawing of a brown apple) for the target descriptions (*red* and *red apple*), there was even a decrease in typicality ratings from the simple description with the adjective alone (*red*) to the complex one (*red apple*)—in other words, a brown apple was a worse example of *red apple* than of *red*, according to Smith and Osherson's (1984) participants. The fact that overextension for adjective-noun combinations depends on the diagnosticity of the particular adjective reflects our color typicality knowledge about the noun categories (apples and canaries).

1.2 *Further Effects of World Knowledge Involving Color Terms*

Later developments employing different methodologies such as corpus/dictionary analysis and color labeling experiments shed further light on the role of world knowledge on color-term usage. Steinvall (2002) analyzed color-adjective uses in the Bank of English corpus and color-adjective entries in the Oxford English Dictionary (1993), and made an important distinction between the ‘classifying’ and the ‘characterizing/descriptive’ functions of color adjectives. While the latter use focuses on the color of a specific instance in the referential setting, the former use, which Steinvall (2002) also called *type modification*, picks out a subtype of the noun category in question. For example, natural kinds (e.g., onions, undyed hair) usually exist only in a few colors, rather than in the full range of a color spectrum, and for these, Steinvall (2002) observed that basic color terms are predominantly used to classify the subtypes based on colors without necessarily being descriptively precise with regard to the actual referent object (e.g., *red onion* for a purple hue, since there are no other types of onions closer to a prototypical bright red). These results are consistent with Anishchanka, Speelman and Geeraerts’s (2014) results from an analysis of color-term usage in online marketing, in which the authors found hypernymous usage with a broad referential range of colors for basic color terms, and a much narrower referential range for non-basic color terms. The exact causal history of basic color terms—whether they arose due to the limited color types in frequently mentioned natural categories, or they just happened to be conveniently adequate for type classification—is, however, admittedly unclear (Steinvall 2002).

Aside from adjective-noun combinations, there are experimental studies involving single-word color terms in color labeling and categorization tasks that also demonstrate an effect of typicality biases due to the object category being asked about. Using hand-drawn images of typically orange (e.g., a carrot) and typically yellow objects (e.g., a banana), Mitterer and de Ruiter (2008) demonstrated that participants were more likely to label, for example, a carrot as “orange” and a banana as “yellow” even when these were presented in the same exact hue. A color typicality effect was also evident at the level of perception for the hue midway between orange and yellow: Participants who saw this ambiguous orange-yellow hue on a carrot first categorized the same color sock (i.e., an object with little intrinsic color bias) in a later task as “orange,” and those who saw this hue on a banana first categorized the same color sock in a later task as “yellow”.

In order to locate the cause of the color typicality effect precisely by teasing visual and declarative memory from life experience apart, Mitterer et al. (2009) picked traffic lights as their visual stimuli in a later experiment. Due to European Union regulations, EU citizens presumably share a common perceptual experience with regard to traffic lights, but some language groups nevertheless differ in their naming of the color of the middle light. Dutch speakers call the middle light *oranje* (‘orange’), whereas German speakers call it *gelb* (‘yellow’). Mitterer et al. (2009)

found that this difference in color-naming habits led to different color-labeling behavior in the experiment when the same ambiguous orange-yellow hue was presented on a traffic light image: When presented with the same exact hue on a middle traffic light, Dutch speakers were more likely to call it “orange,” while German speakers were more likely to call it “yellow,” reflecting their habit. These two language groups were, however, indistinguishable when the ambiguous hue was presented on other object categories, such as a carrot, a banana, and a sock, for which there is presumably no systematic difference in color-naming habits between the two groups. Mitterer et al. (2009) thus concluded that it is primarily declarative memory (i.e., everyday color-term usage), rather than visual memory from life experience, that gave rise to the world knowledge effect in color labeling and categorization.

In sum, there is reason to believe that our general knowledge of typical properties of and relations between objects plays an important role in using color adjectives for labeling different hues. In our study, we investigated whether our knowledge of typical object properties and relations in the world influences our discrete categorization judgments in the context of adjective-noun combinations as well. Specifically, we were interested in the competing factors of set intersection in concept composition and color typicality knowledge in complex concepts with a bias toward a nonfocal color (cf., Heider 1972; Regier et al. 2005, for the notion of a ‘focal’ color, which refers to the best representative of a color category, as widely recognized across different linguistic communities), such as *red hair* (whose typical red is not the focal, bright red) and *green tomato* (whose typical green is much lighter than the focal green). In a series of pilot studies, we first found a set of object categories whose typical colors are not focal colors in our commonsense knowledge (e.g., *red hair* and *green tomato*). Most previous studies involving color descriptions and visual stimuli presented a single image at a time on a given trial, in which the participant had to make a yes-no response with regard to a color description or a forced choice between two color descriptions for the better match. In our Experiment 1, we had participants make a forced choice between two images, one in a focal color and the other in a nonfocal color, for the better match to an adjective-noun description such as *green tomato*. We predicted that, compared to categories without a color typicality bias (e.g., boxes), categories with a typical nonfocal color bias in the real world such as tomatoes would lead to a higher proportion of responses toward the nonfocal-colored image, against the predictions of stricter accounts of concept composition which might accept only good examples of *green* as good examples of *green tomato*. For example, if participants treat the word meanings of *green* and *tomato* separately first and simply combine them in set intersection for the meaning of *green tomato*, they might prefer a focal-green tomato even though it looks artificial. Forced-choice preference data in Experiment 1 by itself, however, would not establish that people’s discrete categorization judgments with regard to an adjective-noun combination differ depending on the intrinsic color properties of the relevant object category. A simple *preference* for a focal-green box image over a nonfocal-green box image as an example of the description *green box* tells us nothing about whether the participant would

categorize the dispreferred nonfocal image (or even the focal image) as an example of *green box* or not in yes-no format. In order to get at people's truth-judgments directly, in Experiment 2 we conducted another picture-phrase matching experiment in which only one image was presented at a time, and participants judged whether the image was an example of the target adjective-noun description in yes-no format. A world knowledge effect in such discrete judgments would strongly point to flexibility in our truth-evaluations even in non-figurative language, contrary to some traditional assumptions in theoretical accounts of color adjective meanings as intersective (e.g., Keenan and Faltz 1985; Chierchia and McConnell-Ginet 1996; Drašković et al. 2013).

2 Experiments

2.1 *Pretest: Category Confirmation and Color Shift Judgments Along a Spectrum*

We presented photographs of seven categories with an intrinsic color bias: banana, bear, jeans, tomato, egg, grass, and horse. All of these categories had at least two naturally existing typical colors, with one of the colors being more 'canonical' (e.g., yellow bananas and green bananas). In order to investigate the effect of a color typicality bias on color judgments on a fine-grained level, we found digital photographs of objects from these categories, and manipulated the color of each category by starting with the original image and creating a duplicate layer with varying levels of transparency, hue, and/or saturation with a color copied from another object image of the same category in Photoshop.¹ The images varied in color along an 11-level spectrum (similar to Hampton 1996), but we needed to make sure that the color manipulation did not affect the category status—e.g., a banana with an ambiguous yellow-green color is judged to be atypical but nevertheless a banana—as observed in categorization judgments and response times.

¹There are two main ways of color matching we used in Photoshop. First, Photoshop has a built-in function 'Match Color' under Image—Adjustments for copying the color of (a selected area in) a source image to (a selected area in) a target image directly. Although this automatic function sometimes affects the luminance patterns in the target image too much, it is possible to control the luminance to preserve the patterns in a precise way. The second way is to create a duplicate layer in the target image file and use an eyedropper to sample a color in a source image to copy onto the target image or simply use the above method and adjust the transparency of the duplicate layer. It is also possible to pre-process the source image by applying an 'Average' filter under Filter—Blur when the internal pattern of the source image is too complex and difficult to transfer naturally to another image using the automatic method. Using the 'Color' mode instead of the default 'Normal' mode in the duplicate layer helps preserve the internal patterns based on contours, etc. Unfortunately there is no single optimal way for all categories, so it is necessary to decide the optimal combination on a case-by-case basis.

In Pretest (a), we looked at three kinds of stimuli: (1) the seven color-biased categories in three different shades (Levels 3, 6, and 9 on the 11-level spectrum, with higher numbers indicating higher color typicality)—a total of 21 Main trials—for detecting any extreme unnaturalness in any direction of the spectrum (3: less typical, 6: midpoint, 9: typical); (2) six Control trials involving categories that require positive extension of the normal noun meaning (stone lion, rubber duck, wooden toy car, model train, Mickey Mouse, and Miffy (an animated rabbit character)), which were included to ensure that our atypical colors in the Main stimuli would not lead to as much surprise as in these noun extension cases; and (3) 24 “No” Filler trials that required a clear “no” response, in order to prevent a set response (see Fig. 1).

Thirteen native speakers of Dutch provided picture-word match decisions in yes-no format under a 5-s time limit, and the response times were measured as well. The judgments were generally consistent with our expectation: For our Main stimuli, we observed around 95% “Yes” responses to all three shades (Levels 3, 6, and 9) of our test categories (258 out of 273 trials, with 1 timed-out trial and 14 “No” responses), confirming that our color manipulation in our stimulus images did not affect their noun category membership—black bears are just as good as brown and black-brown ambiguous ones for the category *bear*, and for Fillers, accuracy was high at 86%. For the six Control categories, for which we expected much greater surprise compared to the atypical colors in the Main trials at least in participants’ reaction times and possibly also in their higher rejection rates, our participants gave 94% “Yes” responses (73 out of 78), accepting the images most of the time for a broader sense of each noun category. In response times, however, these Control categories led to the slowest decisions, as we expected (mean = 1.28 s, see Table 1). Among our Main categories, in contrast, reaction-time differences due to color levels in trials with “yes” responses ($n = 258$: Level 3 average = 1065 ms, Level 6 average = 1009 ms, Level 9 average = 974 ms) were



Fig. 1 Example trials from Pretest (a): (1) Main trial, (2) Control trial, and (3) “No” Filler trial (translated from Dutch). [A color reproduction of this figure appears in the eBook at doi:[10.1007/978-3-319-45977-6_5](https://doi.org/10.1007/978-3-319-45977-6_5)]

Table 1 Response times in noun category confirmation

Stimulus type	<i>N</i>	Mean (ms)	Standard deviation (ms)
Main, Level 3	91	1077 (1065 in “yes” trials)	448
Main, Level 6	91	1073 (1009)	582
Main, Level 9	91	1002 (974)	421
Control	78	1283	589
Fillers	312	1169	518



Fig. 2 An example trial from Pretest 2(b): tomato. [A color reproduction of this figure appears in the eBook at doi:[10.1007/978-3-319-45977-6_5](https://doi.org/10.1007/978-3-319-45977-6_5)]

Table 2 Average level of perceived color shift on a level 1–11 manipulation spectrum

Category	<i>N</i>	Mean	Standard deviation
Banana	13	4.92	1.19
Bear	13	6.08	1.38
Egg	13	5.69	1.65
Grass	13	5.54	1.56
Horse	13	5.54	1.85
Jeans	13	4.85	1.28
Tomato	13	5.08	1.38

small and not statistically reliable ($F(2, 255) = 0.82, n.s.$), confirming our expectation that an atypical color would not make participants hesitate on the category membership of the object shown.

In Pretest (b), we asked the same group of participants from Pretest (a) for their color-shift judgment along a color spectrum, in order to confirm that they did indeed see a color change in our stimuli, and that the locus of this change was not skewed too much toward one end of our color manipulation spectrum. In this task, the participants saw the entire spectrum of 11 colors of each object category on a single screen and indicated the manipulation level at which they thought there was a color shift by typing in the corresponding number (see Fig. 2).

The direction of the spectrum on the screen (from Level 1 to Level 11, or from Level 11 to Level 1) was randomized for each trial. With Level 6 being the mid-point on the scale of 1–11, participants reported a perceived color shift around an average level of 5–7 for all our Main categories, as we expected (see Table 2).

3 Experiment 1: Forced Choice Between a Focal Color Versus a Nonfocal, Typical Color

In order to test which color people choose between a focal color and a nonfocal but canonical color for the category (e.g., focal green vs. nonfocal, ‘tomato’ green) as the better example of an adjective-noun description (e.g., *green tomato*), we conducted a preference judgment task in which participants had to choose between an image pair. As a control, we also tested categories with no strongly associated color (e.g., box). We predicted that participants’ color typicality knowledge would

influence their color preferences in this task, such that participants will be much more likely to prefer a nonfocal color over a focal one for color-biased categories such as tomatoes than for color-neutral categories such as boxes.

4 Method

We gave 11 adult native Dutch speakers a forced-choice picture-phrase matching task in which they saw two photographs of an object along with an adjective-noun combination and picked the image they preferred as the better match for the phrase. For example, the participant would see on the computer screen a photograph of a green tomato in its typical color ('nonfocal' green), another photograph of the same green tomato whose color was manipulated to be a focal green, and the expression *green tomato* (see Fig. 3). 'Nonfocal' colors were simply sampled from the web in a search of photographs of our stimulus categories, and for an operational definition of 'focal' colors in our digital images, we used the RGB triplets in Table 3.

We picked a category from the pretests which had a nonfocal color as a typically existing color and added more to the list for four color-biased categories (green tomato, green apple, orange sky, red leaf) and, as a control, four color-neutral categories (box, flag, table, T-shirt) whose colors were matched with a color-biased category. There were 16 filler trials with non-color adjective modifiers (such as striped apple, bald man, female scientist, and wooden spoon) or with mismatching noun categories. Four of these filler trials had an image of the canonical color for the category (e.g., red tomato) along with a focal-color image (focal-green tomato) to check that participants actually paid attention to the description (*green tomato*)



Fig. 3 Example trials from Experiment 1: (1) color-biased, and (2) color-neutral categories, in which the figure on the *right* of each pair is a focal *green*, and the figure on the *left* is a non-focal *green* typical of a *green* tomato. [A color reproduction of this figure appears in the eBook at doi:10.1007/978-3-319-45977-6_5]

Table 3 RGB triplets for the colors in Experiments 1 and 2

Color name (based on X11 names)	RGB
Blue	(0, 0, 255)
Dark orange	(255, 140, 0)
Office green	(0, 128, 0)
Red	(255, 0, 0)

Table 4 Mean proportion (%) of focal preferences by category in Experiment 1

Biased		Neutral	
Green apple	45	Green flag	55
Green tomato	45	Green box	73
Orange sky	73	Orange table	100
Red leaf	27	Red T-shirt	27

and pick the focal-color as the better match even when it is an unnatural image, rather than simply choose a more familiar image (red tomato, which is a bad match for the description *green tomato*) on most trials. Participants clicked arrow keys to indicate ‘left image,’ ‘right image,’ or ‘no preference’ (the last option was expected for the noun-mismatching trials).

4.1 Results

We predicted a color bias effect in a specific direction, namely, a lower proportion of focal preferences for color-biased categories. We thus recoded the participants’ responses for a one-tailed logistic regression test: Preference for the focal-color image was ‘1,’ and preference for the nonfocal-color image or no preference was ‘0.’ We analyzed the proportion of focal preferences as a function of color bias (see Table 4). Logistic regression with a random slope and intercept for Participant revealed that participants showed a significantly higher proportion of focal preferences for color-neutral categories (mean = 0.64, $SD = 0.487$, $N = 44$) than for color-biased categories (mean = 0.48, $SD = 0.505$, $N = 44$) ($z = -1.72$, $p = 0.043$, one-tailed).

5 Experiment 2: Yes-No Categorization Judgment

Although Experiment 1 demonstrated that, given a pair of images with a focal and a nonfocal color, participants’ preferences for an image matching a target phrase showed an effect of the typical color of the target category, it does not establish that people’s truth-judgments or categorization judgments may differ for the same color on two different objects, depending on whether the object categories have an intrinsically typical color or not. In other words, we were interested in finding instances of a nonfocal color that normally falls outside an acceptable range of a certain color term (e.g., a gingery-orange color for the term *red*) to see if this color will be rejected when it is applied to a category without an intrinsic bias in favor of that color (e.g., a gingery-orange car as an example of *red car*), but accepted when it is applied to a category with a color-biased category (e.g., gingery-orange hair as an example of *red hair*). We thus conducted another picture-phrase matching experiment in which only one image was presented at a time, and participants

judged whether the image was an example of the target adjective-noun combination in yes-no format. In addition to color adjectives, we included some other kinds of adjectives, such as pattern and material adjectives, in order to explore the generalizability of a world knowledge effect.

5.1 Method

For our linguistic stimuli, we selected seven adjectives, each combined with two noun categories to be tested for their ‘Biased’ versus ‘Neutral’ status in the relevant adjective dimension (color, pattern, or material) in a pretest: *red hair/car*, *green tomato/chair*, *striped apple/T-shirt*, *straight leg/road*, *cork mug/board*, *wooden bike/frame*, and *woolen shoe/floor-mat*. We conducted a pretest with 17 native speakers of Dutch to establish the Biased versus Neutral distinction in each of the seven category pairs above using three tasks. For color and pattern adjectives (*red*, *green*, *striped*, *straight*), we first conducted a focal preference task similar to Experiment 1 to confirm a higher focal preference for the neutral noun categories (see Table 5).

For material adjectives (*cork*, *wooden*, *woolen*), Biased versus Neutral status of noun categories was confirmed in free-response production and yes-no typicality judgment tasks. In the free response task, participants were shown a category name (e.g., *bike*) and asked to type in the typical material it is made of. Next, in the typicality judgment task, participants were shown an adjective-noun combination (e.g., *wooden bike*), and asked whether it was a typical combination. Our goal in the free-response task was to find no instances of spontaneous production of our target materials (cork, wood, wool) in the Biased categories, but a few instances of the target (or synonymous/hyponymous) materials in the Neutral categories, and our three category pairs confirmed our expected pattern (see Table 6). In the typicality judgments, we also confirmed the expected pattern of higher positive responses to our Neutral adjective-noun combinations than to the Biased counterparts in our category pairs (see Table 7).

For the 14 adjective-noun combinations, we prepared two photographs for each adjective-noun combination, one focal (e.g., red hair with a bright focal red) and one nonfocal (red hair with a more typical orange/copper hue). Within a Biased-Neutral pair, the values on the relevant adjective dimensions were held constant (RGB for color, pixel proportions for source material, and pattern/shape

Table 5 Mean proportion of focal preferences by category pair in Experiment 2 pretest

Biased		Neutral	
Hair (target ADJ: red)	24	Car (target ADJ: red)	94
Tomato (target ADJ: green)	41	Chair (target ADJ: green)	59
Apple (target ADJ: striped)	0	T-shirt (target ADJ: striped)	88
Leg (target ADJ: straight)	41	Road (target ADJ: straight)	71

Table 6 Common responses in Experiment 2 pretest (free response)

Biased		Neutral	
Mug	13 (out of 17) 'porcelain/stone/clay'	Board	10 'cork'
	no 'cork'		4 'cardboard/paper'
Bike	16 'metal/steel/iron/aluminum'	Frame	14 'wood'
	no 'wood'		
Shoe	17 'leather'	Floor-mat	11 'fabric/cotton'
	no 'wool'		1 'wool'

Table 7 Proportion of 'yes' responses (%) in Experiment 2 pretest (typicality judgment)

Biased		Neutral	
Cork mug	0	Cork board	94
Wooden bike	6	Wooden frame	100
Woolen shoe	6	Woolen floor-mat	59

for pattern by copying and pasting, see Fig. 4). Twenty-four adult native speakers of Dutch saw a photograph along with an adjective-noun combination, and judged (yes/no) whether the picture matched the expression within a three-second time limit. Each participant saw both photographs for each of the 14 adjective-noun combinations for a total of 28 main trials, along with 28 filler trials.

5.2 Results

In the Focal condition, acceptability judgments were high (>80%) for all adjective-noun combinations except one (*striped T-shirt*, 33%), confirming that participants treated the task as a category- or truth-judgment and not just a typicality/familiarity judgment. In the critical Nonfocal condition, in contrast, Biased categories (*hair*, *bike*, etc.) led to significantly higher 'Yes' responses (55%) than Neutral categories (*car*, *frame*, etc., 27%; $p < 0.001$).

5.3 Discussion

Our finding demonstrates that when typical properties of (noun) categories in our commonsense knowledge are biased *against* the 'focal' value of an adjective dimension (e.g., focal red in hair, 100% wood throughout a bike, etc.), our standards for categorization are relaxed such that a 'nonfocal' value (orange/copper rather than red, or wood only in parts of a bike) is more acceptable for these categories compared to those that have no such typicality bias against a focal value. Experiment 2 suggests that similar effects of typicality knowledge play a role in different

Biased	Focal	Nonfocal	Neutral	Focal	Nonfocal
red hair	 87	 92	red car	 100	 17
green tomato	 96	 58	green chair	 96	 92
striped apple	 91	 63	striped T-shirt	 33	 4
straight leg	 96	 38	straight road	 100	 0
cork mug	 87	 46	cork board	 91	 14
wooden bike	 96	 71	wooden frame	 100	 42
woolen shoe	 87	 17	woolen floor-mat	 83	 17

Fig. 4 Stimuli and percentage of “Yes” categorization responses from Experiment 2: non-focal and focal versions of color-biased categories (*left half*) and color-neutral categories (*right half*) (translated from Dutch). [A color reproduction of this figure appears in the eBook at doi:[10.1007/978-3-319-45977-6_5](https://doi.org/10.1007/978-3-319-45977-6_5)]

domains of adjectival meanings, such as colors, pattern, and material, although future research is needed for a much wider range of stimuli. The typicality effect in rapid discrete categorization beyond typicality ratings (e.g., Smith and Osherson 1984) lends support to theoretical accounts that propose a uniform underlying representational space for both typicality and truth judgments, such as Hampton’s (2007) threshold model.

6 Conclusion

Our results point to noun context effects on the interpretation of color adjectives, whose meanings show shifting boundaries for truth-judgments. Color adjectives may seem context-independent and intersective for many categories when the category-specific color spaces converge, but when we consider color judgments for categories with an intrinsic color bias in the real world, we observe context-dependent truth-judgments in uses of color terms. Similar effects of extensional feedback (Hampton 1988) in truth and categorization judgments may arise for many other adjective classes that have traditionally been analyzed as intersective.² Compositional processes that go beyond classical logic and set theory (such as Boolean conjunction and set intersection) are so pervasive in natural language that they cannot simply be set aside as a peripheral issue in semantic theory and pose a serious challenge to accounts of meaning composition as set intersection (e.g., Chierchia and McConnell-Ginet 1996; Heim and Kratzer 1998). It would also be important in future research to pursue further the compositional processes at varying degrees of frequent and conventionalized adjective-noun combinations. Dynamic, context-dependent ‘recalibrations’ of a predicate meaning (Kamp and Partee 1995) or modification of a comparison class to which a predicate applies (Klein 1980) seem to point to general processes of meaning composition in any domain where we have extensional feedback based on our world knowledge, not limited just to a small class of vague predicates.

Experimental studies in categorization and reasoning have made strides in mapping our conceptual space (Gärdenfors 2000) and fine-tuning our ideas about the combination of different conceptual dimensions. Fine-grained quantitative comparisons in the degree or amount of overextension between our study and earlier ones, especially Hampton (1996), would be difficult due to the subtlety of color space and color manipulation (or the lack of detailed descriptions of stimuli in Smith and Osherson 1984). A combination of tasks, such as truth-value judgments, color shift judgments in simultaneous presentation of multiple colors, forced-choice preferences between given colors, and phrase-picture matching tasks, should take us closer to better understanding of concept combination involving color adjectives.

There are important additional insights and challenges from the theoretical literature on color-term interpretation. One is the source or ontological status of the colors—i.e., whether they represent two distinct kinds (e.g., brown vs. black horses) or two stages of the same kind (green vs. red tomatoes that ripen over time). Kennedy and McNally (2010) argue that color adjectives are ambiguous between a gradable reading (denoting a degree scale for the color quality/quantity) and a non-gradable one (denoting a binary presence/absence of an underlying property

²In spontaneous production data (Sedivy 2003), color adjectives do differ, however, from material or scalar adjectives in that they appeared frequently even when they were over-informative and unnecessary, perhaps for a reinforcing effect in referential communication based on the perceptual salience of colors.

correlated with the surface color—e.g., genetic makeup). It remains to be seen whether such taxonomic knowledge is automatically and rapidly accessed, and makes a qualitative difference in our semantic composition of color adjectives with nouns. One may also apply the insights from an account of gradable adjectives such as Toledo and Sassoon's (2011) by analyzing the context-dependent determination of truth-conditions in terms of comparison classes consisting of other members of the same category (in type classification of color-biased categories), or apply the theoretical distinction between stage-level versus individual-level predicates (Carlson 1977) to color adjectives by considering other possible instantiations of an individual (for more gradable usage, e.g., in a maturational sense).

Contrary to a domain such as colors, there are domains that do not have a reasonable context-independent focal point or 'most typical value' (such as size and height—*big, tall*). We would expect similar world knowledge effects for these predicates as well (e.g., a man who is 190-cm tall may be considered tall in normal business attire but not tall in basketball gear, showing the typicality bias in height for basketball players as opposed to height-neutral businessmen), but these predicates need to be studied in future experimental research. Another interesting issue for future research is whether the relative order of modifier and head has any impact on the composition of meanings in real time. It would be interesting to see if preference for nonfocal color typicality is facilitated in languages with post-nominal adjectives, such as French and Hebrew, in which one processes the relevant noun category before a color adjective, in ways that are observable through nonfocal preference speed and proportion measures, compared to Dutch or English, with pre-nominal adjectives.

Acknowledgements This work was supported by the NWO VICI project *Between Logic and Common Sense* while the author was a postdoctoral researcher at the Utrecht Institute of Linguistics OTS. I thank Carien Westenbroek, Nina Wiedenhoff, Eva Poortman, and Yoand Winter for discussion and for help with data collection and stimulus and task selection. I also thank the other project team members and the audience at my Leiden-Utrecht Semantic Happenings talk for valuable comments.

References

- Anishchanka, A., Speelman, D., & Geeraerts, D. (2014). Referential meaning in basic and non-basic color terms. In W. Anderson, C. P. Biggam, C. Hough & C. Kay (Eds.), *Colour studies: A broad spectrum* (pp. 323–338). John Benjamins Publishing Company.
- Cherlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Los Angeles: University of California Press.
- Carlson, G. N. (1977). A unified analysis of the English bare plural. *Linguistics and Philosophy*, 1 (3), 413–457.
- Chierchia, G., & McConnell-Ginet, S. (1996). *Meaning and grammar: An introduction to semantics*. Cambridge, MA: MIT Press.
- Drašković, I., Pustejovsky, J., & Schreuder, R. (2013). Adjective-noun combinations and the generative lexicon. In J. Pustejovsky, P. Bouillon, H. Isahara, K. Kanzaki, & C. Lee (Eds.), *Advances in Generative Lexicon Theory* (pp. 181–202). Dordrecht: Springer.

- Gärdenfors, P. (2000). *Conceptual spaces*. Cambridge: MIT Press.
- Hampton, J. A. (1988). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *14*, 12–32.
- Hampton, J. A. (1996). Conjunctions of visually-based categories: Overextension and compensation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*, 378–396.
- Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science*, *31*, 355–383.
- Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, *93*(1), 10–20.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford, UK: Blackwell Publishers.
- Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, *57*, 129–191.
- Keenan, E. L., & Faltz, L. M. (1985). *Boolean semantics for natural language*. Dordrecht: Kluwer.
- Kennedy, C., & McNally, L. (2010). Color, context, and compositionality. *Synthese*, *174*, 79–98.
- Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, *4*(1), 1–45.
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*, 158–190.
- Mitterer, H., Horschig, J. M., Müsseler, J., & Majid, A. (2009). The influence of memory on perception: It's not what things look like, it's what you call them. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1557–1562.
- Mitterer, H., & de Ruiter, J. P. (2008). Recalibrating color categories using world knowledge. *Psychological Science*, *19*, 629–634.
- Regier, T., Kay, P., & Cook, R. S. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(23), 8386–8391.
- Sedivy, J. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, *32*(1), 3–23.
- Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, *8*, 337–361.
- Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, *12*, 485–527.
- Steinvall, A. (2002). *English colour terms in context*. Doctoral dissertation, Umeå University.
- The Oxford English Dictionary*. (1993). CD-ROM version. Oxford: Oxford University Press.
- Toledo, A., & Sassoon, G. (2011). Absolute vs. relative adjectives—Variance within vs. between individuals. In *Proceedings of Semantics and Linguistic Theory (SALT 21)*, (pp. 135–154).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Concept Typicality and the Interpretation of Plural Predicate Conjunction

Eva B. Poortman

Abstract This chapter studies the interpretations of plural sentences with conjoined predicates, e.g. *The boys are sitting and cooking* and *The boys are waving and smiling*. Such sentences are sometimes interpreted intersectively, sometimes non-intersectively (or ‘split’), and sometimes both interpretations appear to be allowed. This is surprising, since the logical structure of these sentences is identical, i.e. they differ only with respect to content words (e.g. *sitting, cooking* vs. *waving, smiling*). I propose that the logical interpretation of these sentences is systematically affected by lexical information tied to the complex predicate in the sentences, specifically their so-called typicality effects. With a set of experiments, I show that (a) the acceptability of a sentence in a non-intersective situation can be expressed in terms of a continuum and (b) each acceptability proportion is predicted by the typicality of the two conjoined predicates applying simultaneously. This way, I specify at least one of the relevant pragmatic considerations that determine the interpretation of a plural sentence with conjunctive predicates. More generally, these results stress the importance of conceptual structure of predicates in semantic theories of language.

1 Introduction

For a long time, it has been a common tradition in logical semantics to draw a clear line between lexical knowledge and compositional operations (Cruse 1986; Jacobson 2014). Logical semanticists generally focus on the latter, studying the way meaning is composed and how syntax affects this process, mostly irrespective of word meanings. Even though lexical semantics is acknowledged to some degree, the connection between the two domains of study has always been relatively weak. Such an approach proved sufficient in the study of the meanings of logical expressions such as *all* or *at most three*. Whether we speak of *all girls*, *all giraffes*

E.B. Poortman (✉)

Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, The Netherlands
e-mail: evapoortman@gmail.com

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_6

139

or even *all giraffes owned by girls*, does not change the logical contribution of *all* to the meaning of the phrase. One would expect a similar independence of word meaning with logical operators such as *each other* and *and*, e.g. the logical contribution of *and* does not change between *I saw boys and girls* and *I saw giraffes and elephants*. Recently however, studies that looked into these operators revealed that we can no longer make do with such a simple division of labor (e.g. Dalrymple et al. 1998; Kerem et al. 2009; Poortman et al. 2017).¹ These works have shown that contextual information (containing lexical information) and compositional semantics can interact, exposing the need for a more complex story than what is generally assumed. The current chapter deals with one area in which we see ‘logical’ sentence meaning being affected by ‘non-logical’ word meaning, namely plural predicate conjunction, proving that the connection between the two domains is actually fruitful.

Plural sentences with conjunctive predicates as in (1) and (2) are considered to be true if and only if every boy that is referred to is in the intersection of the two sets that are denoted by the conjoined verbs. In other words, sentence (1) is true iff each boy is sitting and each boy is reading, and sentence (2) is true iff each boy is waving and each boy is smiling.

- (1) The boys are sitting and reading
- (2) The boys are waving and smiling

We arrive at such interpretations by applying the well-known boolean analysis of conjunction, according to which it behaves as set-theoretic intersection (Keenan and Faltz 1985; Partee and Rooth 1983), and combining it with a distributivity operator that shifts a VP into one that holds of a plural individual such as *the boys* iff that VP holds of each atomic part of that individual, i.e. each boy (Link 1983). Importantly, such an analysis assumes that the way we logically reason about these natural language sentences is independent of the lexical elements they contain (such as *sitting*, *standing*, *waving* or *smiling*). As a result, the logical interpretations of sentences like (3) and (4) are expected to be derived in a similar way as those of (1) and (2), with the difference between the sentences only being a matter of word meaning.

- (3) The boys are sitting and standing
- (4) The boys are sitting and cooking

In this chapter, I report experiments that show that sentences (3) and (4) in fact receive weaker logical interpretations than sentences (1) and (2). Sentence (3) is generally interpreted such that a subset of the boys is sitting and the rest of the boys are standing—I will call this a ‘split’ interpretation² (Heycock and Zamparelli 2005).

¹In fact, even a seemingly simple term such as *all* has been shown to be context-dependent at times Leslie et al. (2011).

²Throughout the chapter, when using the term ‘interpretation’ I simply refer to the situations that support a truthful usage of a sentence.

Sentence (4) also allows such a ‘split’ interpretation, but crucially to a lesser extent than sentence (3). Understanding such acceptability patterns calls for a systematic investigation of the lexical information that is tied to the verbs in the sentence, as this appears to be inseparable from a proper analysis of conjunction. I show that there is in fact a continuum of acceptability values (i.e. percentages of “true” judgments) for sentences with conjunctive predicates in non-intersective or ‘split’ situations, and I account for this continuum with a principle that predicts how language users apply predicates to plural subjects based on the typicality structure of the complex predicate.

2 Context and Logical Meaning

The insight that context in general, and lexical information in particular, systematically predicts the logical interpretation of a sentence is relatively new. A turning point came with an influential paper on reciprocity by Dalrymple et al. (1998). Dalrymple et al. were the first to start to incorporate a notion of context within logical meaning. They put forward a principle called the Strongest Meaning Hypothesis (SMH), which aims to resolve ambiguity that is caused by contextual information, specifically in the area of reciprocals. Structurally similar reciprocal sentences as in (5) and (6) receive different logical interpretations, despite the fact that they merely differ with respect to the lexical information in the context of the reciprocal expression.

(5) The boys know each other

(6) The boys are following each other

For each occurrence of the reciprocal, the SMH selects as its interpretation the strongest meaning (from an inventory of six possible meanings) that is consistent with context. For example, if we assume three boys, then sentence (5) receives a strong interpretation in which every boy knows every other boy, since there are no contextual restrictions on the amount of possible ‘knowing-relations’. All weaker meanings are consequently disallowed for (5). By contrast, sentence (6) most likely means something weaker than every boy following every other boy. According to the SMH, we weaken the meaning of the sentence as far as context pushes us to. For this example, that meaning is most likely one where boy 1 follows boy 2, and boy 2 follows boy 3. This is the strongest candidate meaning that does not contradict our knowledge about following people (assuming the boys are not following each other in a circle). Summarizing, we can say that the SMH, unlike what logical semanticists would assume, takes non-logical information to be relevant in determining logical interpretation. It does so without delving too much into what exactly constitutes this non-logical information, other than referring to it as ‘context’ in general. However, without a specific notion of context, empirically supporting such a principle is very difficult.

Nevertheless, Winter (2001) re-uses the main gist of the SMH as a solution to the different interpretations of plural predicate conjunction. Many previous works have described non-intersective interpretations of plural sentences such as example (3) above, as well as sentences with noun phrase conjunction as in (7) given below (e.g. Krifka 1990; Heycock and Zamparelli 2005).

(7) John and Mary wrote an article together

Sentence (7) does not entail that John wrote an article together and Mary wrote an article together, similar to sentence (3) which does not entail that the boys are sitting and those same boys are standing. Krifka (1990) proposes to extend the generally accepted non-intersective conjunction of noun phrases as in (7) to conjunction of predicates as in (3). He proposes that any conjunction P_1 and P_2 holds of an entity x if x can be partitioned into two entities x_1 and x_2 such that P_1 holds of x_1 and P_2 holds of x_2 . For (3), this means that whenever the entity ‘the boys’ can be partitioned into two entities, then the predicate *sitting* can hold of one of these entities and the predicate *standing* can hold of the other. Winter (2001) acknowledges that while this is a proper analysis for sentences like (3), it fails to capture the fact that sentences like (1) only allow an intersective interpretation. He claims that on top of Krifka’s descriptive proposal of non-intersective conjunction in addition to intersective conjunction, we also need a principle that determines when which analysis is allowed, thus when the different interpretations actually occur. Winter (2001) proposes that a maximality principle like the SMH is a suitable candidate. First, he assumes that the SMH is not construction-specific to plural sentences with reciprocals. He rephrases it into a general principle of plural predication, such that any complex plural predicate with a meaning that is derived from one or more singular predicates using universal quantification is interpreted using the logically strongest truth conditions that are not contradicted by known properties of the singular predicate(s) (Winter 2001). Note that unlike Dalrymple et al.’s SMH, Winter’s extended SMH does not speak of ‘context’ in general, but focuses on a more manageable part of context, namely the lexical information tied to predicates. The contrast between minimal pairs like (1) and (3) is then captured in the following way. Again, the SMH selects the logically strongest possible candidate meaning for each sentence. When a strong interpretation (intersective conjunction) is consistent with properties of the predicates, then this is the attested meaning of the sentence—an example is sentence (1). On the other hand, when such a strong interpretation is inconsistent with these properties, the interpretation is weakened. We see this in sentence (3): An intersective interpretation in which all boys are in the intersection of the set of sitting individuals and the set of standing individuals contradicts what we know about ‘sitting’ and ‘standing’. Thus, sentence (3) receives a ‘split’ interpretation, which is the strongest interpretation that does not contradict this knowledge.

In the current chapter, I argue that the predictions made by the SMH can be too strong, specifically because its notion of context is still not defined specifically

enough. Consider sentences (8) and (9), which are of a similar nature to sentence (4) above.

- (8) The men are lying down and drinking
- (9) The men are waving and drawing

If Winter's extended SMH is correct in assuming that non-intersective interpretations are only available when intersective interpretations are strictly ruled out by the predicates, then these sentences would only allow an intersective interpretation. As I will show in the current chapter, non-intersective, 'split' interpretations are readily available to many speakers for sentences like (4), (8) and (9), even though the predicates do not strictly exclude an intersective one. For example, it may be exceptional, but it is possible for a person to wave and draw simultaneously.

Several previous works have recognized a similar problem for the SMH concerning reciprocal sentences that receive weaker interpretations than predicted (e.g. Winter 2001; Philip, 2000; Kerem et al. 2009; Poortman et al. 2017). For example, both Kerem et al. (2009) and Poortman et al. (2017) showed experimentally that a sentence like *The boys are pinching each other* in the case of three boys is judged as true in a situation where each boy pinches only one other boy, despite the fact that a stronger interpretation is not excluded by properties of the predicate *pinch*.

3 Typicality: Defining Context

These examples, both with reciprocals and conjunction, point to a fundamental issue with the proposal at hand. Since context is not specified in much detail, the SMH, both in its original and extended form, assumes that the interpretation of these sentences is only sensitive to so called 'definitional' aspects of the meaning of predicate concepts. In other words, it only takes into account whether particular denotations of predicates are possible or impossible, i.e. whether they are an instance of that predicate concept or not. In the case of predicate conjunction, that means that the hypothesis only looks at whether intersective conjunction is possible or not, given the predicates at hand. Such sharp distinctions appear to be insufficient in accounting for the interpretation patterns that we observe. Alternatively, one can take into account so-called typicality effects in categorization. The notion of typicality (at least as it is assumed in the current study) simply refers to the phenomenon that human subjects are able to grade different instances of a concept with respect to their representativeness of a given category. To illustrate, besides being able to categorize a *sparrow* and an *ostrich* within the *bird* category and a *bat* and a *crocodile* outside of it, people also distinguish between members of a category: e.g. a *sparrow* is judged a more typical bird than an *ostrich*. Since the 1970's, a range of psychological studies has shown for such one-place predicates that subjects consistently rank some instances of a concept as more typical than others, and that such rankings correlate with other measures of typicality such as categorization speed

and error rate (e.g. Rosch 1973; Smith et al. 1974; Rosch and Mervis 1975). Moreover, I follow Hampton (2007) in assuming that an instance's category membership and its typicality as an instance of that category are two related behavioral measurements, based on one and the same underlying variable. For example, there is a correlation between binary membership measures for *sparrow* (1), *ostrich* (1), *bat* (0) and *crocodile* (0) on the one hand, and their typicality rating on the other hand (*sparrow* > *ostrich* > *bat* > *crocodile*). Hampton assumes a so-called threshold model, according to which there is a threshold somewhere along a typicality function that makes a binary distinction between members (*sparrow*, *ostrich*) and non-members (*bat*, *crocodile*).

For the current purposes, taking into account typicality means extending the aspects of meaning that the SMH is sensitive to from definitional to prototypical, thus fleshing out what constitutes context.³ Incorporating such typicality effects on reasoning was first proposed as a solution for reciprocal sentences, in the shape of the Maximal Typicality Hypothesis (Kerem et al. 2009; Poortman et al. 2017). Firstly, this hypothesis assumes that typicality effects also exist for verb concepts like the binary predicate concept *pinch*, i.e. it assumes that subjects can consistently rank some instances of pinching as more typical than others. The difference with noun concepts like *bird* is thus merely a matter of the *type* of things that are being categorized, namely events instead of objects. Secondly, it predicts that these typicality effects for verb concepts systematically affect the logical interpretation of the reciprocal expression that they combine with. Specifically, the Maximal Typicality Hypothesis (MTH) predicts the core situation for a reciprocal sentence to be the maximal one among those that are most typical for the predicate concept in the sentence (for an elaborate discussion see Poortman et al. 2017).

In this chapter, I extend the same logic to plural sentences with predicate conjunction. I view the MTH as a general principle of meaning composition that systematically governs vagueness in plural sentences. The MTH as such a general mechanism surfaces whenever a graded concept such as reciprocity (in reciprocal sentences) or distributivity (in predicate conjunction sentences)⁴ combines with a natural concept such as a verb or an adjective—which each have their own typicality structure. Accordingly, I claim that typicality also affects interpretation in

³Note that Dalrymple et al.'s notion of context is not very clear, and seems to include the predicate in the scope of the reciprocal as well as things like world knowledge and speaker intentions. All I mean here is that I study one particular aspect of what they refer to as 'context', namely the predicate concepts, and I use typicality as a probe into it.

⁴I assume that reciprocity and distributivity are graded similar to how simple plural sentences are (Winter 2017). Take for example the sentence *The men are sitting*. Such a sentence is more often judged true the more men are actually sitting. Similarly for reciprocal sentences like *The men know each other* (which is more often judged true the more knowing pairs there are) and predicate conjunction sentences like *The men are sitting and cooking* (which is more often judged true the more men are both sitting and cooking).

plural sentences where two predicate concepts are conjoined. Note that my aim here is not to provide a theory of concepts nor to explain how typicality judgments come about (for an explanation of relevant notions of typicality see Hampton (2017)), but merely to investigate the relationship between interpretation and typicality. The proposal works as follows.

We know that predicate conjunctions such as (1) through (4) are classically analyzed intersectively, such that both conjoined predicates apply to each individual in the plural subject simultaneously. The MTH predicts that the degree to which a weaker interpretation is available depends on typicality. Take for example the predicate concept *sitting*, for which we assume typicality effects much like for the concept *bird*: within the instances that are categorized as ‘sitting’ instances (or members of the *sitting* category), I predict that some are consistently judged more typical than others. For example, an event in which a person is sitting straight up in a chair is probably judged as a more typical instance of *sitting* than an event in which a person is leaning so far back that they are almost lying down on the floor. I expect that the different interpretation patterns of sentences (1)–(4) arise due to similar typicality effects with verb concepts. To illustrate, I predict that an event in which a person is sitting while also reading can easily be categorized as a typical instance of the concept *sitting*, and similarly an event in which a person is reading while also sitting can be categorized as a typical instance of the concept *reading*. The fact that both predicates apply simultaneously (i.e. a person is reading while sitting, or sitting while reading) does not affect the typicality of the event for each predicate concept in isolation. For predicate combinations for which this is the case, I predict that plural sentences with combined predicates (e.g. *the boys are sitting and reading* in example (1)) simply behave according to an intersective analysis, i.e. we multiply the number of times the two predicates apply simultaneously based on the number of individuals that the plural refers to, since this does not affect the typicality of the entire situation. By contrast, I predict that events in which a person is sitting while also standing are physically impossible, and therefore not categorized as instances of *sitting*.⁵ Similarly, they are not instances of *standing* either. Assuming a threshold model (Hampton 2007), we could say that they fall below the threshold for category membership of *sitting* or *standing* (when the typicality of an event increases for the concept *sitting*, it decreases for the concept *standing*, and vice versa). If we now consider a plural sentence with these predicates combined (e.g. *the boys are sitting and standing* in example (3)), an intersective interpretation is physically impossible (or of close-to-zero typicality for each concept), causing us to weaken the interpretation such that each individual that the plural subject refers to either sits or stands. Crucially, the MTH does not merely make predictions about these two extreme cases, but in fact considers them to be end points on a scale. An example like *the boys are sitting and cooking* (example (4)) clarifies this. Consider

⁵Or, if you could imagine some strange situation of sitting and standing simultaneously, then it would at least be a highly atypical instance of *sitting*. This does not affect the nature of the argument.

an event in which a person is sitting while also cooking. Such an event, however odd, can probably be categorized as an instance of the concept *sitting* (i.e. it is above the category membership threshold). And, similarly, an event in which a person is cooking while also sitting can be categorized as an instance of the concept *cooking*. Interestingly however, the fact that both predicates apply at the same time, causes the typicality of the event as an instance of each concept in isolation to decline. An event in which one is sitting while cooking is probably not the most typical instance of *sitting*, and an event in which one is cooking while sitting is definitely not a typical instance of *cooking*.

Crucially, this degree of atypicality that I predict is caused by the simultaneous application of two predicates within one event. Therefore, I will henceforth speak about ‘compatibility’ as a measurement of this atypicality, allowing me to directly compare compatibility between different pairs of predicates. Ideally, if one were to be interested in the full typicality structure of verb concepts like *sitting*, one could construct a standard task (similar to the tasks used for noun concepts like *bird*), namely rating all possible *sitting* events with respect to how typical they are for the concept. In the current chapter however, I restricted my measurements since they are led by a direct research question, based on an observation: I was specifically interested in the different interpretation patterns of sentences like (1)-(4), thus looking for a measurement that allowed me to directly compare “sitting and cooking” versus “sitting and reading” versus “sitting and standing”. One should keep in mind, however, that when I speak of ‘compatibility’ of predicate concepts P_1 and P_2 , I aim to indirectly measure the typicality of an instance of P_1 in an event that has been categorized as an instance of P_2 , and the typicality of an instance of P_2 in an event that has been categorized as an instance of P_1 . The reason I did not test this directly, i.e. presenting subjects with an event in which both predicates apply and measure its typicality as an instance of concept P_1 and of concept P_2 (similar to measuring the typicality of an *ostrich* as an instance of the concept *bird*), was because I was also interested in strictly incompatible pairs of predicates (e.g. *sitting and standing*)—which cannot be depicted within one event. Moreover, I did not use a direct textual test either, i.e. “rate how typical it is for a person to do P_1 in a situation in which she is known to be doing P_2 ” because this seemed to me a harder and more confusing task than a simple compatibility task. Thus instead, I conducted a more indirect, simple textual compatibility test, in which I assess the typicality of an event for concept P_1 and concept P_2 by measuring the predicates’ compatibility.

The measured compatibility is predicted to affect the way sentences are interpreted in which those predicates combine with a plural. Specifically, the less typical the intersective situation is for the two combined predicate concepts in isolation, the more we diverge from an intersective interpretation when those combine with a plural. In more general terms, one could say that when we interpret sentences like (1) through (4) there are two factors at work: (1) maximize the number of predications and (2) retain typicality. These factors are sometimes in conflict, which is when the MTH surfaces. Summarizing, I phrase the proposal as follows:

MTH for plural predicate conjunction

For any sentence *The X P₁ and P₂*, in which *X* is a plural and *P₁* and *P₂* are singular predicates; and an event *E* in which *P₁* and *P₂* apply simultaneously to one individual member of *X*:

The less typical event *E* is for concept *P₁* and for concept *P₂*, the more we diverge from an intersective interpretation of the sentence *The X P₁ and P₂*.

Crucially, this formulation of the MTH assumes that both the notions of typicality and of acceptability can be expressed in terms of a continuum—allowing for more subtle distinctions than the SMH. The experiments that are discussed below measure typicality (via compatibility) and interpretation separately. I predict to find a) that there is a continuum of typicality values for event *E* as an instance of predicate *P₁* and of predicate *P₂*, b) that there is a continuum of acceptability values for a plural sentence with those predicates (*The X P₁ and P₂*) in a given situation and c) that the values on both continuums correlate—indicating that typicality of an event for particular concepts in isolation systematically affects interpretation of sentences containing those concepts. I conducted two behavioral experiments and a correlation analysis to test these predictions.

4 Experimental Investigation

This section reports on pretests, two experiments and a correlation analysis. Experiment 1 checked the acceptability of plural predicate conjunction sentences of the form *The X are P₁ and P₂* (where *X* is a plural noun and *P₁* and *P₂* are predicates) in a non-intersective, ‘split’ situation. Experiment 2 measured compatibility of predicate concepts *P₁* and *P₂* as an indirect typicality test, as argued in the previous section. Materials for the experiments were constructed based on pretests that were conducted in order to include a wide range of compatibility values in the actual experiments.

4.1 Pretests: Constructing Materials

The aim of the first pretest was to gather as many Dutch verb combinations as possible, especially atypical ones. I provided 8 participants with 16 sets of two pairs of predicates, *P₁* and *P₂* and *P₁* and *P₃*: one very natural pair, and one pair that is physically impossible to apply simultaneously, e.g. *sitting and reading* (*P₁* and *P₂*) and *sitting and standing* (*P₁* and *P₃*). I then asked them to provide as many verbs that they could come up with that combine with *P₁* (i.e. *sitting* in this case) that led to a possible but atypical, uncommon or strange pair. The pairs that participants constructed, combined with more natural pairs that I came up with, led to a list of 91 verb combinations in total.

In the second pretest, 29 different participants rated all of these 91 pairs for compatibility, in a paper-and-pencil task. For each pair, participants were asked to rate how odd⁶ they would consider it if both verbs applied to one person at the same time. Oddness was rated on a 6-point scale, where 1 meant ‘not odd at all’ and 6 meant ‘physically impossible’. I mentioned explicitly that 5 thus meant ‘very odd, but physically possible’, in order to distinguish large atypicality from impossibility, or in other words: to distinguish members from non-members by indicating that the category membership threshold for P_1 or P_2 is between 5 and 6. Results of this pretest showed great variability in ratings between verb pairs, with a high level of agreement between the participants (Cronbach’s alpha was 0.88 for the 91 items). The selection of verb pairs that were to be used in Experiments 1 and 2 proceeded as follows. I defined sets of verb pairs on the basis of the different P_1 verbs, e.g. a set consisted of *sitting and reading*, *sitting and standing*, *sitting and knitting*, *sitting and cooking*, etc. Then I selected the 12 sets that showed the greatest range of ratings. Finally, three verb pairs from within each of these 12 sets were selected: the verb pair that was rated lowest on the oddness scale (compatible pairs like *sitting and reading*), the verb pair that was rated highest (incompatible ones like *sitting and standing*), and a verb pair that was rated in between, at a mean of 4 points⁷ (atypical pairs like *sitting and cooking*). The 36 verb pairs that constituted the final material, translated from Dutch, are given in Table 1 (the original Dutch material can be found in the Appendix). Creating the three groups (with labels ‘compatible’, ‘incompatible’ and ‘atypical’) was done purely to ensure variability while constructing the materials. I will refer to these three groups when discussing set-up and results of Experiments 1 and 2. Note however that the distinction between the groups is not meaningful in the final correlation analysis of all data points.

4.2 *Experiment 1: Interpretation of Plural Predicate Conjunction Sentences*

This experiment checked the acceptability of 36 plural sentences with two conjoined verbs in a ‘split’ situation. Each sentence was of the form *The X are P_1 and*

⁶Phrasing the question negatively by asking ‘how odd’ subjects would rate a situation was done because (a) directly asking for ‘how compatible’ they would judge two predicates seemed like a too technical and too direct task, and (b) asking for ‘how typical’ they would judge a situation turned out to be ambiguous in Dutch. Some subjects interpreted the word *typical* to mean ‘atypical’, whereas asking for oddness is unambiguous.

⁷Additional inclusion criteria included that (a) each verb should be expressed by one word only, (b) ratings for verb pairs should have small variation (whenever there was more than one candidate for selection, the one with the lowest standard deviation for the ratings was selected). Finally, if after considering these criteria there were still two candidate pairs for the atypical group, I decided that (c) atypical verb pairs should have no 6 point ratings (since that meant that at least one participant judged it to be physically impossible for the two verbs to apply simultaneously). This was only a very small criterion, applying to one case.

Table 1 Overview of predicate pairs, translated from Dutch

Compatible	Incompatible	Atypical
Sitting and reading	Sitting and standing	Sitting and cooking
Waving and smiling	Waving and clapping	Waving and drawing
Walking and singing	Walking and swimming	Walking and writing
Crawling and screaming	Crawling and jumping	Crawling and reading
Standing and reading	Standing and squatting	Standing and falling asleep
Reading and smiling	Reading and sleeping	Reading and drawing
Lying down and stretching	Lying down and running	Lying down and drinking
Drawing and yawning	Drawing and typing	Drawing and walking
Swimming and smiling	Swimming and crawling	Swimming and reading
Texting and frowning	Texting and knitting	Texting and waving
Knitting and singing	Knitting and clapping	Knitting and walking
Sleeping and drooling	Sleeping and telephoning	Sleeping and standing

P_2 (where X is a plural noun and P_1 and P_2 are verbal predicates). The reason for using a ‘split’ situation was that sentences with incompatible pairs cannot be depicted any other way, and I wished to keep all factors in the comparison between pairs equal.

Participants A total of 33 students from Utrecht University (28 female, age $M = 21$) participated for monetary compensation. All participants were native speakers of Dutch without dyslexia. Prior to the experiment all participants signed an informed consent form.

Materials The material consisted of two versions of a truth-value judgment task, each containing 18 unique test items plus 18 filler items that were the same across versions. Each test item contained a plural predicate conjunction sentence in Dutch (*The X are P_1 and P_2*)⁸ and a drawing depicting four individuals in a non-intersective, ‘split’ interpretation of that sentence: predicate P_1 applied only to persons 1 and 2, predicate P_2 applied only to persons 3 and 4. Half of the pictures depicted male individuals, and the other half depicted female individuals. An example of a test item drawing is given in Fig. 1.

In each version of the experiment, one third of the test items contained sentences with verb pairs that were considered compatible P_1 and P_2 in the second pretest (e.g. *The men are sitting and reading*), one third contained sentences with verb pairs that were considered incompatible P_1 and P_2 (e.g. *The men are sitting and standing*) and one third contained sentences with pairs that were considered atypical P_1 and P_2 (e.g. *The men are sitting and cooking*). The same drawings were used for sentences with compatible and incompatible pairs with identical P_1 (e.g. *The men*

⁸All the sentences in the experiment were in the simple present tense, which can be used to describe ongoing events as well as states in Dutch. Whereas in English one would use the progressive tense for all sentences in Experiment 1, the distribution of the progressive tense in Dutch is different, such that it could not be used for all sentences in the experiment alike.



Fig. 1 Example of a test item drawing of Experiment 1. All rights for further use of the illustration arranged with the artist

are sitting and standing and *The men are sitting and reading*). To ensure that subjects never saw the same drawing twice (such as the one in Fig. 1), one of these sentences occurred in version 1 and the other occurred in version 2. The atypical items were divided over the two versions, resulting in two experiments with 6 sentences with compatible pairs, 6 sentences with incompatible pairs and 6 sentences with atypical pairs each, accompanied by 18 unique drawings.

Filler items contained similar drawings with four people, but a different type of accompanying sentence. The accompanying sentences in the filler items were either sentences with quantifiers (*Some boys are P*) or sentences mentioning specific individuals in the picture (*Boys A, B and C are P*). Half of the filler items were expected to be judged true, and half of them were expected to be judged false. Both versions of the experiment contained the same filler items.

The order of items was pseudo-randomized using Mix software (Van Casteren and Davis 2006), with the following restrictions: items containing the same verb were at least six items apart; there were at most two test items immediately following each other, and at most two filler items immediately following each other; similar test items (in terms of compatible/incompatible/atypical) or similar filler items (in terms of quantifier/specific individuals) never immediately followed each other. Finally, I constructed two orders of each version, with the second one having reversed order of items.

Procedure Each participant completed one version of the experiment. The task was presented in a sound-proof booth on a PC using Open Sesame software (Mathôt et al. 2012). Prior to entering the sound-proof booth, each participant received verbal instructions explaining the experimental set-up. Further, more detailed instructions were given on the PC monitor.

After being instructed, each participant completed three practice trials. Subsequently, they were given the opportunity to ask for clarifications, if necessary.

No verb used in the practice session appeared in the actual experiment. The experiment itself consisted of the 36 items described above. Drawing and sentence were presented in the center of a white screen. Participants were instructed to indicate as soon as possible whether they judged the sentence to be true or false given the situation in the drawing by pressing the left or right button with their dominant hand.

Coding and analysis Responses were coded ‘1’ when participants judged a sentence to be true for a given drawing, and ‘0’ when they judged a sentence to be false. I computed the proportion of true-responses for each of the three types of sentences for each participant. I then performed a repeated measures ANOVA across participants with Compatibility as the within-subjects factor (with 3 levels: compatible, atypical, and incompatible).⁹ Post hoc Bonferroni corrected multiple comparisons were performed in order to analyze differences between different Compatibility levels in detail. An ANOVA across items, with Compatibility as the between-item variable (also with 3 levels), gave similar results to the participant analysis. Therefore only the first analysis is reported.

Results Table 2 provides an overview of the data. It shows the acceptability of sentences, i.e. the percentage of “true” judgments, for the three levels of Compatibility that were tested for all versions taken together. More detailed results on acceptability per item are in the Appendix. Overall, the truth percentages of the different sentences in the experiment ranged from 24% to 100%. I predicted lowest acceptability for the sentences with compatible pairs and highest acceptability for the sentences with incompatible pairs.

A repeated measures ANOVA revealed that there was a main effect of Compatibility ($F(1.36, 43.49) = 37.41, p < 0.001$). This means that the mean proportions of acceptability for the three Compatibility levels are not equal. Pairwise comparisons show that all three levels differ significantly from each other in acceptability: the acceptability of sentences with compatible predicates differs from the acceptability of sentences with incompatible predicates ($p < 0.001$); the acceptability of sentences with compatible predicates differs from the acceptability of sentences with atypical predicates ($p < 0.001$); and the acceptability of sentences with incompatible predicates differs from the acceptability of sentences with atypical predicates ($p < 0.05$). Note again, however, that the main conclusion from

Table 2 Mean acceptability

Compatibility type	% “true” judgments (st. dev.)
Compatible	54 (31)
Atypical	78 (31)
Incompatible	84 (32)

⁹A repeated measures ANOVA with Version as between-subjects factor was also performed, but showed no effect of Version ($F(3, 29) = 0.47, p = 0.71$) nor an interaction effect of Version * Compatibility ($F(6, 58) = 0.82, p = 0.58$). I thus collapsed the versions for the analysis.

this experiment is not that there are significant differences between groups, but the fact that there is variation in the data.

4.3 Experiment 2: Compatibility of Predicate Pairs

This experiment checked compatibility for the 36 predicate concept pairs that were used in sentences of Experiment 1. I aimed to measure the typicality of one particular event as an instance of the concepts P_1 and P_2 , namely one in which both predicates P_1 and P_2 apply simultaneously (event E). As already discussed at length earlier on in this paper, I conducted an indirect textual test in which I assess the typicality of event E for P_1 and P_2 by measuring the predicates' compatibility. This test was identical to the pretest, but carried out by different subjects and now containing fewer items, in a fully controlled experiment.

Participants The same 33 students from Utrecht University from Experiment 1 participated in this experiment. Each subject completed the interpretation experiment first, before proceeding with the typicality experiment. Also, in between experiments they took part in a third, unrelated experiment.

Materials The materials consisted of a questionnaire containing 36 statements about one person involved in two actions simultaneously. Half of the statements were about males and half of them were about females (matching the gender of persons in the pictures of Experiment 1). Each statement contained a singular object (a male or a female) and two conjoined predicates (e.g. *The man is sitting and reading*). The 36 pairs of verbs were the same as the ones used in sentences of Experiment 1, thus one third of the pairs were considered compatible in the second pretest (e.g. *sitting and reading*), one third were considered incompatible (e.g. *sitting and standing*), and one third were considered atypical (e.g. *sitting and cooking*).

The order of items was pseudo-randomized using Mix software (Van Casteren and Davis 2006), with the restriction that at most two items of the same type (in terms of compatible/incompatible/atypical) immediately followed each other.

Finally, four different orders of the questionnaire were constructed: two versions that started with the statements about males (with the second one having reversed order within males and females statements), and two versions that started with the statements about females (with the second one having reversed order within males and females statements).

Procedure Each participant received one of the questionnaires on paper, in a sound-proof booth. They were instructed to rate how odd¹⁰ they would consider it if both verbs applied to the given person at the same time. Oddness was rated on a

¹⁰As mentioned in footnote 6, phrasing the question negatively by asking 'how odd' subjects would rate a situation was done because directly asking for 'how compatible' or 'how typical' they would judge a situation turned out to be unsuitable.

6-point scale, where 1 meant ‘not odd at all’ and 6 meant ‘physically impossible’, or in other words: to distinguish members from non-members by indicating that the category membership threshold for P_1 or P_2 is between 5 and 6. It was mentioned explicitly that 5 thus meant ‘very odd, but physically possible’, in order to distinguish large atypicality from impossibility.

Coding and analysis Responses were coded ‘1’ through ‘6’ corresponding to the participant’s oddness judgment. This way the incompatibility rating for each verb pair was computed. I performed a repeated measures ANOVA with Compatibility as the within-subjects factor (with 3 levels: compatible, atypical, and incompatible). Post hoc Bonferroni corrected multiple comparisons were performed in order to analyze differences between different Compatibility levels in detail.

Results Table 3 provides an overview of the data. It shows the mean incompatibility rating for the three levels of Compatibility that were tested, for all versions taken together. More detailed results on incompatibility rating per item are in the Appendix. Overall, mean ratings per verb pair ranged from 1.03 to 5.94, and there was a very high correlation between these ratings and the ratings for these items in the pretest ($r = 0.98, p < 0.001$).

A repeated measures ANOVA revealed that there was again a main effect of Compatibility ($F(1.95, 62.45) = 1187.02, p < 0.001$). This means that the mean incompatibility ratings for the three Compatibility levels are not equal. Pairwise comparisons show that all three levels differ significantly from each other: the incompatibility of supposed compatible pairs differs from the incompatibility of supposed incompatible pairs ($p < 0.001$); similarly for the incompatibility of compatible vs. atypical pairs ($p < 0.001$); and the incompatibility of incompatible vs. atypical pairs ($p < 0.001$). This means that the three groups that were selected based on the pretest were confirmed in Experiment 2 (with different subjects and a subset of the stimuli).

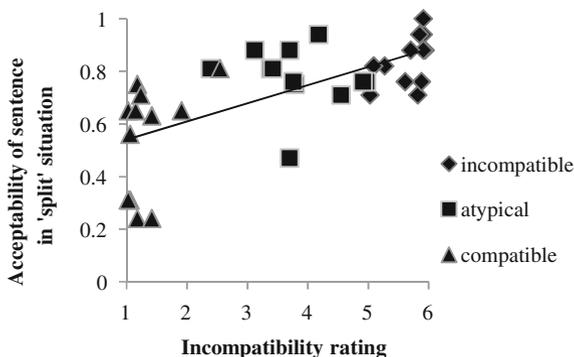
4.4 Correlation Between Interpretation and Compatibility

The crucial test for the proposal is the relationship between interpretation and compatibility. In order to account for the degree to which non-intersective interpretations of sentences *The X are P_1 and P_2* are available given two particular conjoined predicates P_1 and P_2 , we need to check whether this correlates with the degree to which P_1 and P_2 are incompatible (as an indirect measurement of the atypicality of P_1 and P_2 applying simultaneously (event E) for each concept in

Table 3 Mean incompatibility ratings

Compatibility type	Mean incompatibility (st. dv.)
Compatible	1.35 (0.29)
Atypical	3.82 (0.58)
Incompatible	5.66 (0.38)

Fig. 2 Relation between mean acceptability (in proportions) of sentences in a ‘split’ situation and mean incompatibility rating of predicate pairs.



isolation). In order to check this, I performed a correlation analysis between all the results of Experiment 1 and those of Experiment 2 (Fig. 2). The result was a positive correlation between mean proportion acceptability of a sentence in a non-intersective interpretation and mean incompatibility rating of a predicate pair ($r = 0.66$, $n = 36$, $p < 0.001$).

5 Discussion

This paper reports on an experimental investigation into the interpretation of plural sentences with predicate conjunction, and its connection to typicality. I proposed that the extent to which non-intersective interpretations are available directly correlates with the atypicality of an event in which the two predicates apply simultaneously. Experiment 1 revealed a continuum of acceptability values of 36 sentences in a non-intersective, ‘split’ situation, ranging from 24% to 100% acceptable. Such a continuum is unexpected under the extended SMH by Winter (2001), which assumes that any given sentence is either true or false in a particular situation, depending on what the context allows. Next, Experiment 2 showed that differences in compatibility exist between different predicate pairs. The compatibility ratings for 36 pairs ranged over the entire 6-point scale. I assumed that the compatibility measurement is an indirect measurement of typicality, namely of the typicality of event E (in which two predicates apply simultaneously) for each predicate concept in isolation, and hence that this effect is similar to the effects that were found repeatedly for one-place predicates (e.g. Rosch 1973). I proposed to extend the Maximal Typicality Hypothesis (Kerem et al. 2009; Poortman et al. 2017) by formulating it for predicate conjunction in such a way that typicality relates to acceptability so that the less compatible the two predicates in Experiment 2 are judged to be (i.e. the less typical event E is), the more a non-intersective interpretation is available. Based on a correlation analysis, I conclude that this prediction was borne out. Note that this correlation does not hinge on my assumption that compatibility is an indirect way of measuring typicality. I merely

take these results to be an indication of conceptual structure of predicates playing a crucial role in sentence interpretation, in line with similar results on reciprocal sentences (Poortman et al. 2017).

5.1 Reference Shift of the Plural Subject?

The particular interpretation that was in the focus of the current study was the so-called ‘split’ interpretation in which P_1 always applies to two of the individuals in the picture and P_2 always applies to the two remaining individuals. I have claimed that this interpretation is sometimes available for predicate conjunction sentences, namely to the degree that a situation in which the conjoined predicates apply simultaneously is atypical. One might argue instead that the acceptability of these sentences given a split interpretation has nothing to do with typicality. As an alternative, one might reason that we accept a sentence like (3) (repeated below as (10)) because its deep structure is the sentential conjunction in (11), which contains two definite plurals that hence allow the possibility of referring to two different groups of boys. In other words, the reasoning would be that we accept (10) in a split situation because we are able to very quickly shift the reference of the plural noun *the boys* from one set of boys to another set of boys. My experiments would then in fact deal with reference resolution instead of with matters of typicality.

(10) The boys are sitting and standing

(11) The boys_i are sitting and the boys_j are standing

If we indeed actually interpret the predicate conjunction in (10) as sentential conjunction (as is made explicit in (11)), then I would expect to see no differences between different test sentences. If reference shift would explain why sentence (10) is accepted in a split situation, then we should be able to use this strategy across the board for all types of predicate conjunction that were tested—whether they are typical, atypical or incompatible. This is clearly not the case, and the question remains what explains the range of acceptability values.

In fact, a pilot study¹¹ has revealed that when sentential conjunction is explicit in the surface form of the sentence (i.e. when subjects are given sentences like (11)), we see that indeed it is possible to shift the referent for different types of predicate

¹¹The study was conducted with 9 participants who were students at Utrecht University (6 female, age $M = 23$), and checked the acceptability of 12 plural sentences with sentential conjunction in a ‘split’ situation. Each sentence was of the form *The x are P_1 and the x are P_2* (where x is a plural noun (used twice) and P_1 and P_2 are verbal predicates). Half of the P_1 and P_2 pairs were compatible predicates while the other half were incompatible predicates (based on pretests from the study reported in this paper). One participant accepted none of the sentences, the remaining eight participants accepted all or all but one.

conjunction. Subjects accepted sentence (11) (with incompatible predicates) given a split situation, but they also accepted sentences like (12) and (13) given a split situation—even though we have seen that sentences containing such compatible predicates behave differently when presented as mere predicate conjunctions (i.e. they are generally judged false 50% of the time). A single subject did not accept sentence (11) in a split situation, but they also refused to accept sentences like (12) and (13) in a split situation.

(12) The boys are sitting and the boys are reading

(13) The boys are waving and the boys are smiling

What this suggests is that reference shift is independent from the obtained results in the current study. Sentences with sentential conjunction do not show the same correlation with typicality as sentences with predicate conjunction do. I believe reference shift of the plural subject in sentences like (11)–(13) is purely motivated by trying to make a sentence true. This explains why most participants always accepted such sentences in a split situation (i.e. used reference shift to make the sentence true). The one subject that did not use reference shift, was consistent in not using it across different types of predicate conjunction (compatible and incompatible).

Another argument against the reference shift explanation is the finding (based on a small pilot¹²) that sentences with proper name conjunctions instead of definite plurals are accepted in a split situation significantly more often when the conjoined predicates are incompatible (as in (14)) compared to when they are compatible (as in (15)).

(14) John, Bill, Sue and Jane are sitting and standing

(15) John, Bill, Sue and Jane are sitting and reading

For such sentences, reference shift of the subject *John, Bill, Sue and Jane* is obviously not possible, and still a sensitivity to the predicate concepts in the sentence is observed, which is along the same lines as the results presented in the current chapter.

Summarizing, I conclude that it is unlikely that the presented results are due to reference shift of the plural subject. Shifting the reference of the subjects from one referent to another does not explain the systematic variability in acceptability, nor that a similar pattern arises for sentences with proper name conjunction as in (14) and (15).

¹²This study was conducted with 22 participants who were students at Utrecht University (18 female, age $M = 19$). It checked the acceptability of 8 plural sentences with predicate conjunction in a ‘split’ situation. Each sentence was of the form *A, B, C and D are P_1 and P_2* (where *A, B, C* and *D* are names and P_1 and P_2 are verbal predicates). Half of the P_1 and P_2 pairs were compatible predicates while the other half were incompatible predicates (based on introspection). Sentences with compatible pairs were accepted in a split situation 10% of the time, sentences with incompatible predicates were accepted in a split situation 40% of the time.

5.2 *Other Measures of Typicality*

Despite the fact that we can safely rule out reference shift as an alternative explanation of the results, obviously there are many other factors that are worth further exploration. The correlation that was found in this study was high ($r = 0.66$, $n = 36$, $p < 0.001$), though obviously not perfect. This means that there must be more factors that affect interpretation besides the one tested here. An important next step is to delve deeper into typicality effects for complex predicates. In the current chapter, I report an experiment that indirectly measured one particular typicality measurement with one particular dependent measure, namely the typicality of two simultaneous actions, rated on a scale. One can imagine that in fact the typicality of the opposite event, i.e. two predicates applying to two separate individuals, or perhaps sequentially to one individual, might also affect the interpretation of a plural sentence with those predicates. Moreover, as pointed out by a reviewer, perhaps not only the verb concepts but also the head noun of the sentences play a role. It might be that the compatibility of two predicates is quite different in the context of humans than it is for example in the context of dogs: people can run and scratch their heads simultaneously, but dogs cannot. In order to fully understand the factors that influence sentence interpretation, an intricate combination of typicality measures is necessary.

Also, it will be good to correlate rating measures with different kinds of dependent measures such as categorization speed or error rate to have a more robust result—similar to the investigations into typicality effects for nouns. However, the fact that even one measure can distinguish different types of verb pairs so clearly, is a promising starting point for this enterprise.

Another related issue is the deeper question of how typicality effects come about: What exactly makes a particular instance of a concept typical? A potential candidate factor is that typicality is formed by prior experiences or likelihood of a situation. An anonymous reviewer, however, pointed out example (16).

(16) The boys are unicycling and juggling

The reviewer claims that despite the fact that we probably rarely see a person simultaneously unicycling and juggling, we still probably interpret the conjunction in sentence (16) intersectively (though of course a full sample of participants would need to be consulted to be sure). Such an example points out that typicality is not simply a matter of frequency, but a far more complex notion that needs to be studied further. The question of what makes something typical does not affect the results described in this chapter per se, but knowing what affects typicality would give them more explanatory power, as pointed out by this reviewer.

5.3 Further Areas

Another logical step would be to investigate other cases in language where typicality affects reasoning. So far we have seen that understanding both reciprocal sentences and the sentences with conjunction that were investigated in the current chapter, is inseparable from the study of concepts. Another area where we see typicality affecting interpretation, is the area of adjective-noun constructions such as *red hair* (Lee 2017). For such a construction, the typicality structure of *hair* appears to interact with the way we interpret the adjective *red*. Even though the concept *red* in isolation might have as its most typical instance a focal red, orange-like hues are generally more typical for the concept *hair*. When the two combine, these typicality preferences interact (for more on these effects see the work by Lee (2017) and Winter (2017)). This interaction is intuitively of a similar nature to the one between a verb concept like *pinch* and the reciprocal expression *each other*, as well as the one between verb concepts like *sitting* and *cooking* and the logical expression *and*. It is highly likely that these are not the only areas in which this is the case, thus it is worthwhile for further research to investigate whether a principle like the MTH can function as a general principle of language use.

6 Conclusion

This chapter started from the observation that plural sentences with conjunctive predicates do not always receive the same logical interpretations. Previous work on reciprocal sentences has already taught us that lexical information can influence sentence meaning in systematic ways (e.g. Dalrymple et al. 1998; Kerem et al. 2009; Poortman et al. 2017). Here I reported on experimental investigation of plural sentences with predicate conjunction, that provided insight into specifically the role of typicality information of predicate concepts. With this result, I add to the line of work that investigates the interface between lexical and compositional semantics, and lead the way towards directions for further research in this area.

Acknowledgements This work was supported by NWO (The Netherlands Organisation for Scientific Research) VICI grant number 277-80-002, “Between Logic and Common Sense: the Formal Semantics of Words” (2010–16). This chapter formed the basis for Chap. 3 of my doctoral dissertation entitled ‘Concepts and Plural Predication: The Effects of Conceptual Knowledge on the Interpretation of Reciprocal and Conjunctive Plural Constructions’, published by LOT (Utrecht, The Netherlands). I am grateful to Hanna de Vries, Marijn Struiksmā, James Hampton, Martin Everaert, Choongkyu Lee, Sonia Kasyanenko, Chris Barker, two anonymous reviewers of the ESSLLI workshop “Formal and Experimental Pragmatics”, two anonymous reviewers of this volume, and the audiences at the Concept Composition workshop (Utrecht 2013), the NYU

Semantics Group (New York 2014) and the Formal and Experimental Pragmatics workshop (Tübingen 2014) for useful remarks and questions. Special thanks to Yoad Winter for extensive discussions throughout the course of this research. Finally, I am thankful to Nir Kerem for initial ideas that eventually led to the study reported in this paper. The illustration in Fig. 1 was made by Ruth Noy Shapira.

Appendix

This appendix contains the original Dutch stimulus verb pairs P_1 and P_2 that were used in Experiments 1 and 2, and their English translation. In Experiment 1, each verb pair was contained in a sentence of the form *The X are P_1 and P_2* (where X is a plural noun). The two rightmost columns in the appendix give the mean results of each experiment for each stimulus verb pair.

Group	Verb pair P_1 and P_2	English translation	Experiment 1 Interpretation (% true)	Experiment 2 incompatibility (rating 1–6)
incomp	Zitten en staan	Sitting and standing	0.94	5.91
incomp	Zwaaien en klappen	Waving and clapping	0.82	5.27
incomp	Lopen en zwemmen	Walking and swimming	0.88	5.94
incomp	Kruipen en springen	Crawling and jumping	0.71	5.82
incomp	Staan en hurken	Standing and squatting	0.94	5.85
incomp	Lezen en slapen	Reading and sleeping	0.76	5.88
incomp	Liggen en rennen	Lying down and running	0.88	5.91
incomp	Tekenen en typen	Drawing and typing	0.82	5.09
incomp	Zwemmen en kruipen	Swimming and crawling	1	5.91
incomp	Smsen en breien	Texting and knitting	0.71	5.03
incomp	Breien en klappen	Knitting and clapping	0.88	5.7
incomp	Slapen en bellen	Sleeping and telephoning	0.76	5.61
	MEAN		0.84	5.66
atyp	Zitten en koken	Sitting and cooking	0.88	3.12

(continued)

(continued)

Group	Verb pair P_1 and P_2	English translation	Experiment 1 Interpretation (% true)	Experiment 2 incompatibility (rating 1–6)
atyp	Zwaaien en tekenen	Waving and drawing	0.75	3.79
atyp	Lopen en schrijven	Walking and writing	0.81	3.39
atyp	Kruipen en lezen	Crawling and reading	0.71	4.55
atyp	Staan en dutten	Standing and falling asleep	0.47	3.7
atyp	Lezen en tekenen	Reading and drawing	0.94	4.18
atyp	Liggen en drinken	Lying down and drinking	0.81	3.42
atyp	Tekenen en lopen	Drawing and walking	0.76	3.76
atyp	Zwemmen en lezen	Swimming and reading	0.76	4.97
atyp	Smsen en zwaaien	Texting and waving	0.81	2.39
atyp	Breien en lopen	Knitting and walking	0.88	3.7
atyp	Slapen en staan	Sleeping and standing	0.76	4.91
	MEAN		0.78	3.82
comp	Zitten en lezen	Sitting and reading	0.65	1.03
comp	Zwaaien en lachen	Waving and smiling	0.31	1.06
comp	Lopen en zingen	Walking and singing	0.24	1.18
comp	Kruipen en schreeuwen	Crawling and screaming	0.81	2.55
comp	Staan en lezen	Standing and reading	0.65	1.91
comp	Lezen en glimlachen	Reading and smiling	0.31	1.03
comp	Liggen en rekken	Lying down and stretching	0.65	1.15
comp	Tekenen en gapen	Drawing and yawning	0.75	1.18
comp	Zwemmen en lachen	Swimming and smiling	0.24	1.42
comp	Smsen en fronsen	Texting and frowning	0.56	1.06

(continued)

(continued)

Group	Verb pair P_1 and P_2	English translation	Experiment 1 Interpretation (% true)	Experiment 2 incompatibility (rating 1–6)
comp	Breien en zingen	Knitting and singing	0.71	1.24
comp	Slapen en kwijlen	Sleeping and drooling	0.63	1.42
	MEAN		0.54	1.35

References

- van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, 38(4), 584–589.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge, U.K.: Cambridge University Press.
- Dalrymple, M., Kanazawa, M., Kim, Y., Mchombo, S., & Peters, S. (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy*, 21, 159–210.
- Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science*, 31(3), 355–384.
- Hampton, J. A. (2017). Compositionality and concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 95–122). London: Springer.
- Heycock, C., & Zamparelli, R. (2005). Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural Language Semantics*, 13, 201–270.
- Jacobson, P. (2014). *Compositional semantics*. Oxford, U.K.: Oxford University Press.
- Keenan, E., & Faltz, L. (1985). *Boolean semantics for natural language*. Dordrecht: D. Reidel.
- Kerem, N., Friedmann, N., & Winter, Y. (2009). Typicality effects and the logic of reciprocity. In *Proceedings of Semantics and Linguistic Theory, SALT19* (pp. 257–274).
- Krifka, M. (1990). Boolean and non-boolean ‘and’. In L. Kálmán & L. Pólós (Eds.), *Papers from the second symposium of logic and language*. Budapest: Akadémiai Kiadó.
- Lee, C. (2017). Typicality knowledge and the interpretation of adjectives. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 123–138). London: Springer.
- Leslie, S., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, 65, 15–31.
- Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bauerle et al. (Eds.), *Meaning, use and interpretation of language*. Berlin: De Gruyter.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314–324.
- Partee, B., & Rooth, M. (1983). Generalized conjunction and type ambiguity. In R. Bauerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use and interpretation of language*. Berlin: De Gruyter.
- Philip, W. (2000). Adult and child understanding of simple reciprocal sentences. *Language*, 76(1), 1–27.
- Poortman, E. B., Struiksma, M. E., Kerem, N., Friedmann, N., & Winter, Y. (2017). Reciprocal expressions and the Maximal Typicality Hypothesis. To appear in *Glossa*.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.

- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 81(3), 214–241.
- Winter, Y. (2001). Plural predication and the strongest meaning hypothesis. *Journal of Semantics*, 18, 333–365.
- Winter, Y. (2017). Critical Typicality: Truth judgements and compositionality with plurals and other gradable concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp 163–190). London: Springer.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Critical Typicality: Truth Judgements and Compositionality with Plurals and Other Gradable Concepts

Yoad Winter

Abstract Compositional semantic frameworks often compute the extension of a complex expression directly from the extensions of its parts. However, much work in cognitive psychology has shown important challenges for compositional methods. For instance, Hampton (J Exp Psychol Learn Mem Cognit 14(1):12–32, 1988b) showed that speakers may let the complex nominal *sports that are games* include *chess* as one of its instances, without admitting chess in the extension of *sports*. Similarly, Lee (2017) experimentally supports the common intuition that instances of *red hair* are not necessarily categorized as *red*. This paper reviews further results about *plural quantifiers*, showing similar challenges for compositionality. It is proposed that typicality effects play a systematic role in compositional interpretation and the determination of truth-values. For instance, the “overextension” effect in the *red hair* example is predicted by the fact that focal red is an atypical hair color. Similarly, in the plural sentence *the men are walking and writing*, the availability of the split reading (“some men are walking and some men are writing”) increases due to the atypicality of doing both activities at the same time (Poortman, 2017). Further, in reciprocal sentences like *the three men are pinching each other*, the number of pinching acts may be three. This is related to the atypicality of situations where every man pinches two other men at the same time, as required by a strong interpretation of *each other*. The paper gives a uniform account of truth-value judgements on these different constructions, based on the identification of conflicts between typical preferences.

1 Introduction

One of the biggest challenges for semantic theory is the tension between contextuality and compositionality. To take a simple example, suppose that someone tells us that *A and B cost the same*. We may reasonably conclude that if A is expensive, then

Y. Winter (✉)

Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, The Netherlands
e-mail: y.winter@uu.nl

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3, DOI 10.1007/978-3-319-45977-6_7

163

B is expensive as well. However, as a general inference rule, this conclusion is too hasty. We realize that when we consider the possibility that A is a laptop and B is a car. At some level, our initial conclusion was justified: without doubt, price is the most salient parameter in our interpretation of the concept EXPENSIVE. However, as soon as we become aware of further information about the objects A and B, we may easily discharge conclusions that were drawn on the basis of price alone. A theory of reasoning that ignores such contextual information, e.g. by relegating it to some underspecified pragmatic variables, runs the risk of becoming empirically empty.

It is important to note that the same risk threatens formal semantic theories of meaning composition. When analyzing the compositional interpretation of phrases like *my expensive laptop* and *my expensive car*, we again have to address problems about contextuality. Here, the interpretation of the adjective *expensive* is most directly affected by its immediate linguistic context: the head noun—*laptop* or *car*. The effect that these nouns have on the meaning of *expensive* is similar to the contextual effect that the identity of A and B have in the example above. Therefore, a theory of semantic composition must pay close attention to contextual effects, and not only consider the formal description of language structures. In view of this point, some researchers in cognitive psychology doubt that formal semantic theories, with their arcane symbolic methods, have a prominent role to play in the analysis of concept composition (Barsalou, 2017).

However, there is another side to the compositionally problem, which indicates that formal theories cannot so easily be dismissed. Consider the distinction between the phrase *expensive laptops and cars* and the phrase *laptops and expensive cars*. The former expression is ambiguous (it may or may not be about expensive cars) whereas the latter is unambiguous (it must be about expensive cars). The ambiguity in the former phrase is not simply a matter of the larger prosodic distance between the words *expensive* and *cars*. This can be seen by considering the phrase *expensive laptops in cars*, which unambiguously refers to expensive laptops, not to expensive cars. Clearly, the syntactic nature of the construction—a coordinate structure or a subordinated prepositional phrase—has an important impact on potential ambiguities when concepts evoked by language are put together. To understand such effects we need to rely on syntactic theory. This is where notions of compositionality from formal semantics become important. As we move on to more complex quantificational structures, temporal and spatial expressions, anaphoric items etc., we increasingly need to rely on structural analyses and symbolic principles to explain the close interactions between form and meaning (Pelletier, 2017).

One of the motivations for the present work is the conviction that both contextual considerations and formal considerations are important for analyzing meaning composition. More specifically, this paper proposes that compositional theories should rely on *typicality effects*. These effects are critical for understanding concepts in general, and they have been playing an important role in the study of concept composition since the early works of Osherson and Smith in the 1980s. Notwithstanding, there is an on-going debate on whether typicality effects are relevant for the study of natural language semantics (Sassoon 2013). Following previous work, especially Kamp and Partee (1995), this paper aims to examine the role that typicality effects

play in formal semantics. I propose that using observations about typicality effects, we can systematically account for gradability phenomena in truth-value judgements. I claim that the distribution of judgements about “truth” and “falsity” of complex sentences is at least partially predictable from judgements about typical instances for the concepts that sentences evoke. Thus, although typicality effects are logically separate from truth-value judgements, they affect such judgements in a non-trivial way.

One might suppose that the more typical an object is for a category, the likelier it is to be categorized in that category. However, part of the interest in typicality phenomena is that their interactions with truth-value judgements are much subtler than that. For instance, as Lee (2017) shows, ginger-like hues are much more often accepted for *red hair* than they are for *red car*. As we will see, similar effects appear in other areas of semantic theory, specifically reciprocal expressions (‘each other’), and conjunctive plural predicates (‘are big and red’, Poortman, 2017). In all three cases, a gradable concept combines with another concept, with possibly conflicting typicality preferences. Capitalizing on this parallelism, the paper employs one unified principle for explaining recent experimental results in the three domains. Section 2 reviews familiar notions about vagueness of gradable concepts. Section 3 uses these notions for also modeling some vagueness effects with plurals. Section 4 addresses familiar motivations to distinguish vagueness. Section 5 explains the proposed approach in relation to the familiar problem of typicality effects with gradable adjectives (‘expensive car’, ‘red car’). Section 6 applies the same approach to plurals, specifically reciprocal expressions (‘each other’) and conjunctive predicates (‘big and red’). Finally, Sect. 7 puts the pieces together, by analyzing the effects of typicality on truth-value judgements. Section 8 concludes.

2 Vagueness and Acceptability Functions

“Truth” and “falsity” are often seen as the pinnacles of natural language semantics. However, truth-conditional semantics must address the fact that speakers’ truth-value judgements about sentences are often inherently non-uniform, in a way that cannot be explained by ambiguity or by multiple processing strategies. Consider for example a simple sentence like *this hue is red*. For monochromatic light at 610 nm, between typical red and typical orange, speakers may disagree on whether the sentence is ‘true’ or ‘false’, or find it hard to decide between the two possibilities (Bonini et al. 1999). The phenomenon is commonly referred to as *vagueness*.¹ We say that vague sentences such as *this hue is red* may have a varying degree of acceptability depending on the situation. Upon being shown a black hue, virtually all speakers (who are not color-

¹For a reader on vagueness in philosophy, see Keefe and Smith (1996). For linguistic work on vagueness, see Kennedy (2011) for a semantic overview, and Nouwen et al. (2011) for further work in linguistics and philosophy. See Alxatib and Pelletier (2011), Egré et al. (2013), Hampton (1998, 2007) Serchuk et al. (2011) for experimental work.

blind) would judge the sentence as ‘false’. Upon being shown a hue in the focal red area (around 650 nm), virtually all speakers would judge the sentence as ‘true’. For the oranges, pinks, magentas etc., we take the acceptability to be a number between 0 and 1, according to the proportion of speakers who consider the sentence true. Proportion-based analyses of acceptability go back to Black (1937) “consistency profiles”. Black lumps between-speaker consistency together with within-speaker consistency. However, there are empirical distinctions between the two (McCloskey and Glucksberg 1978). I consider here acceptability as a between-speaker measure, because it is in line with the way truth-value judgements and typicality judgements are analyzed in the experiments cited throughout this paper.² However, since this paper concerns the connections between acceptability and typicality, it is only crucial that they are consistently measured. Other measures of acceptability and typicality may be used, and are expected to show similar behaviors to those studied here.

Since concepts like RED, EXPENSIVE, TALL etc. impose graded acceptability on situations, we refer to them as *gradable concepts* (GCs). Different theories of vagueness have different interpretations of what acceptability scores for GCs mean and how they should be treated. Here we only use them for reflecting experimental measures of truth-value judgements, without a commitment to any specific theory of vagueness. For convenience, we also ignore all sorts of questions about the non-linguistic context. Thus, for a sentence like *this hue is red* we may assume that the ‘situation’ is a simple color grid presented without any specific context, ignoring complex problems that may appear when the contextual information is more intricate (e.g. as in Egré et al. 2013). These assumptions are useful for addressing our main problem here, and they do not affect too much the proposed analysis.

Summarizing, we define acceptability functions as follows.³

- (1) Given a gradable concept GC, the **acceptability function** ACC_{GC} is a function from situations to the real numbers in $[0, 1]$.

In many cases, the order that GCs induce on situations is *upward-monotone* in the relevant dimension. To see what this means, consider for instance the following sentence.

- (2) This jacket is expensive.

The concept EXPENSIVE satisfies the following condition:

For any two situations S_1 and S_2 with the same jacket x , if the price of x in S_1 is less or equal to its price in S_2 , then $\text{ACC}_{\text{EXPENSIVE}}(S_1) \leq \text{ACC}_{\text{EXPENSIVE}}(S_2)$.

²Measuring speakers’ judgements embodies an “internalist” approach to meaning (Pelletier, 2017). James Hampton (p.c.) remarks that speakers’ acceptability judgements on expressions like *risk* may diverge quite substantially from “externalist” definitions of their meanings (e.g. based on probability theory). This divergence may reflect two different senses of the word *risk*—an “ordinary” sense and a “scientific” sense. This distinction hardly affects the concepts studied in this paper.

³In semantics, the numeric values that acceptability functions return are often called “degrees” (Kennedy 2007). In works on concepts, degrees are often described in terms of “graded membership” in a category (Hampton 1998, 2007; Kamp and Partee 1995) among others).

Monotonicity implies that the acceptability function for the concept EXPENSIVE induces a partial order on situations on the basis of the price. For instance, in sentence (2), suppose that we consider two situations: S_1 , where the jacket costs \$50, and S_2 , where it costs \$120. We expect sentence (2) to be at least as acceptable in situation S_2 as it is in S_1 . Thus, when speakers are asked out-of-the-blue if sentence (2) is “true”, we expect the proportion of positive responses for situation S_2 to be greater or equal than for situation S_1 . Similarly, for the concept RED, the nearer the hue of an object x is to the definitely red area (say, around 650 nm), the higher the acceptability of the sentence x is red. Note that not all gradable concepts are upward-monotone in this way. For instance, the concept MIDDLE-AGED is not upward-monotone in age: its acceptability may increase up to certain ages (say around 55), but at later ages it may start to decrease (say around 65). Thus, this acceptability is upward-monotone in some age groups, and downward-monotone in others.

3 Distributivity and Reciprocity as Gradable Concepts

Noun phrases like *the children* and *the townspeople* were not traditionally studied as part of the puzzles surrounding vagueness in natural language. However, in some works on plurals it has been observed that such noun phrases also exhibit considerable vagueness.⁴ For instance, consider the following example by Lasersohn (1999).

(3) The townspeople are asleep.

How many of the townspeople need to be asleep in order for sentence (3) to be considered “true”? When trying to answer such questions, we see that sentence (3) has a characteristic behavior of vague sentences.⁵ To capture such vagueness effects, we associate an acceptability function ACC with the predicate “are asleep” in (3). To be precise, we associate this acceptability function with a concept DIST, for a *distribution quantifier*. This means that we roughly interpret sentence (3) as claiming that *many*, or *enough*, of the townspeople are asleep. As with these explicitly vague concepts, the acceptability of sentence (3) gets higher as more townspeople are asleep. This monotonicity is determined by the acceptability function for the DIST concept. Explicitly:

⁴See Brisson (1998), Burnett (2012), Dowty (1987), Lasersohn (1999) for theoretical work, and Križ and Chemla (2015), Schwarz (2013) for recent experimental work.

⁵E.g. consider the vagueness criteria surveyed by Kennedy (2011). Sentence (3) has borderline cases, e.g. when 75% of the townspeople are asleep. The same sentence also demonstrates the “Sorites Paradox”: if you consider (3) true for a situation with 10,000 sleeping townspeople, you’ll find it hard to change your mind if one of them suddenly wakes up. Sentences like (3) have also been demonstrated to be context sensitive. E.g. consider Lasersohn’s (1999) example “the subjects are asleep”, in the context of an experiment where the purpose is that all the subjects fall asleep. In such cases the preference that all subjects are asleep becomes stronger. See Burnett (2012) for further discussion. Some authors refer to vagueness effects as in (3) as ‘non-maximality’ or ‘homogeneity violation’ (Križ and Chemla 2015).

- If each townspeople in S is asleep: $\text{ACC}_{\text{DIST}}(S) = 1$
- If no townspeople in S is asleep: $\text{ACC}_{\text{DIST}}(S) = 0$
- ACC_{DIST} is upward-monotone with respect to containment between sets: if the set of townspeople is the same in situations S_1 and S_2 , and the set of sleeping townspeople in S_1 is contained in the set of sleeping townspeople in S_2 , then $\text{ACC}_{\text{DIST}}(S_1) \leq \text{ACC}_{\text{DIST}}(S_2)$.

Similar vagueness appears with another plurality phenomenon, which is related to distributivity: *reciprocity* (Dalrymple et al. 1998; Langendoen 1978). Consider the following sentence.

(4) The three girls admire each other.

Intuitively, sentence (4) is perfectly true if each girl admires the other two. The less admiration relations there are, the less acceptable the sentence is. In experiments by Kerem et al. (2009) and Poortman et al. (2017), sentences similar to (4) were evaluated by participants in different situations, which are described schematically in Fig. 1.

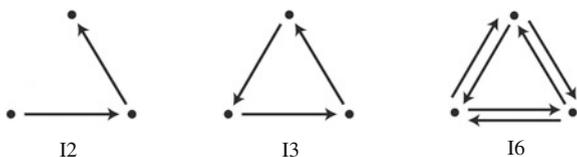
Participants in Poortman et al.’s experiment were presented with schemes as in Fig. 1 and were asked to judge if a given scheme is a “possible depiction of the sentence”. Of the participants, 96%, 48% and 4%, respectively, judged schemes I6, I3 and I2 as possible descriptions for a Dutch translation of sentences like “A, B and C admire each other”. We describe these facts using an acceptability function $\text{ACC}_{\text{RECIP}}$ for the reciprocity concept RECIP:

- In situation I6: $\text{ACC}_{\text{RECIP}}(I6) = 0.96$
- In situation I3: $\text{ACC}_{\text{RECIP}}(I3) = 0.48$
- In situation I2: $\text{ACC}_{\text{RECIP}}(I2) = 0.04$
- The acceptability function $\text{ACC}_{\text{RECIP}}$ is upward-monotone with respect to containment between binary relations. Thus, if the set of three girls is the same in two situations S_1 and S_2 , and the set of admiration pairs in S_1 is contained in the set of admiration pairs in S_2 , then $\text{ACC}_{\text{RECIP}}(S_1) \leq \text{ACC}_{\text{RECIP}}(S_2)$.

The measured acceptability values for situations I2, I3 and I6 are consistent with the monotonicity assumption about $\text{ACC}_{\text{RECIP}}$, since the set of arrows in I6 contains the set of arrows in I3, and the set of arrows in Figure I3 contains the set of arrows in I2.

We have considered three different kinds of gradable concepts (GCs) with adjectives (EXPENSIVE, RED), quantificational distributivity (DIST), and reciprocal quanti-

Fig. 1 Three situations for the interpretation of sentence (4)



fiers (RECIP). As we have seen, all three kinds of GCs are vague and display upward-monotonicity. With this background on GCs, we can address some problems surrounding their interpretation and typicality effects.

4 Acceptability Versus Typicality

Speakers often judge an object to be an atypical member of a concept, but without denying its membership in the concept's extension, or *category* (Rosch 1973). As Kamp and Partee (1995, p. 133) mention, if Robbie is a pelican then his acceptability as an instance of the concept BIRD ought to be close to 1. Notwithstanding, Robbie is clearly an atypical instance of this concept. Since Rosch and her associates' seminal works in the 1970s, this dissociation between acceptability and typicality has been illustrated by a variety of *typicality effects*. For instance, in McCloskey and Glucksberg's (1978) experiments, the following category-exemplar pairs received typicality ratings lower than 7 (on a 1–10 scale), but membership was considered positive by more than 90% of the participants:

ANIMAL: *cobra, lizard, woman* BIRD: *penguin* CARPENTER'S TOOL: *crowbar* DISEASE: *alcoholism, schizophrenia* INSECT: *louse, silkworm* NATURAL EARTH FORMATION: *sinkhole* SCIENCE: *agriculture* WEATHER PHENOMENON: *dew* VEGETABLE: *rhubarb, soy bean*.

This kind of empirical dissociation between measures of acceptability (=graded membership) and measures of typicality has been repeatedly demonstrated, and is hardly controversial. A bigger controversy surrounds the question whether typicality effects relate to the same mental faculties that affect vagueness. Hampton (2007) describes one position in the debate by stating that “Osherson & Smith and Kamp & Partee contend [that] typicality T has quite different properties as a variable from degree of membership M ”. By contrast, Hampton himself (a.o.) proposes “that both degree of membership M and typicality T are based on a single underlying metric of similarity [to the prototype]”. For more on this debate, see Hampton (1998), Rips (1989).

The present paper shows more evidence that typicality affects category membership.⁶ Specifically, typicality for lexical concepts may influence graded membership in categories for complex expressions made of these lexical concepts. For a start, this section considers some examples where typicality is distinguished from acceptability, and which are specifically relevant for the rest of this paper. These examples concern the behavior of noun concepts, reciprocity and distributivity, as tested in recent experiments by Kerem et al. (2009), Lee (2017), Poortman (2017) and

⁶Note however, that I remain agnostic regarding Hampton's claim that typicality and acceptability are determined by the same underlying metrics of similarity (to a prototype). Especially, the proposal below is neutral with respect to Hampton's assumption that typicality measures with *complex expressions* reflect the same metrics that affects truth-value judgements.

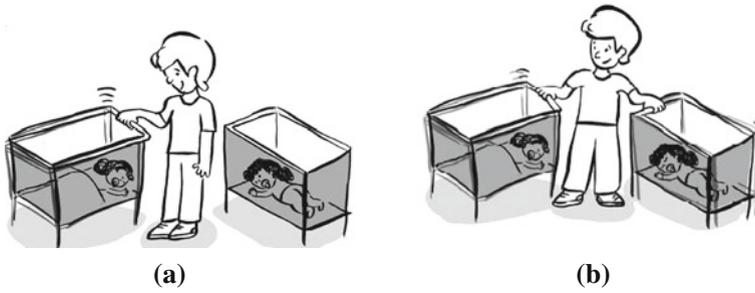


Fig. 2 Two illustrations of the transitive meaning of the Hebrew verb for “shake”. All rights for further use of the illustration arranged with the artist

Poortman et al. (2017). These examples help us observe common principles in the complex interplay between acceptability and typicality.

Example 1 HAIR

In one of Lee’s experiments (Lee, 2017: Fig. 5.4), participants were shown a picture of hair that was atypically dyed focal red. 87% of the participants accepted this picture as an instance of the concept RED HAIR. Therefore, we may reasonably conclude that the stimulus was quite acceptable as an instance of HAIR.⁷ However, in a color preference task for the concept HAIR, only 24% of the participants preferred focal red to a more natural hue for hair (RGB value: 201,113,13). In accordance with naive intuition, we may conclude that hair that is dyed focal red is an acceptable instance of HAIR, but is atypical for the category.

Example 2 SHAKE

In one of Kerem et al.’s experiments, participants were shown two pictures of a man shaking infant beds (Fig. 2a, b), as well as an incomplete transitive sentence “the man is shaking (something)”.⁸ Although both pictures clearly show the man doing the activity reported in the sentence, 87% of the participants preferred Fig. 2a as a “better illustration for the sentence”. We conclude that a situation with two patients is an atypical instance of the concept SHAKE, despite its high acceptability.

Example 3 WALK and WRITE

In one of the experiments by Poortman (2017), participants were instructed to rate the oddness of a situation where a person *walks* and *writes* at the same time.

⁷Lee did not directly check the degree of acceptability for HAIR. Overextension effects (Hampton 1996) may mean that the actual acceptability for HAIR in the focal red stimulus was slightly below 87%. But this effect cannot be too big, since the pictures that Lee used were quite natural hair.

⁸The Hebrew example in Kerem et al.’s questionnaires was *ha-iš mena’ane’a* (“the-man shake”). The Hebrew verb only has the transitive meaning of the English verb *shake* (“cause to move”), not the intransitive meaning (“tremble”).

The average oddness rate for this pair of verbs was 3.39 on a scale between 1 and 6, where 1 meant “not odd at all” and 6 meant “physically impossible”. We conclude that such a situation is atypical for the concepts WALK and WRITE. Nevertheless, the situation is highly acceptable as a legitimate instance for each of those concepts: when Lucy happens to be walking and writing at the same time, we do not expect many speakers to reject the description “Lucy is walking” just because the event is atypical for the concept WALK. Similarly we do not expect the sentence “Lucy is writing” to be rejected if she happens to be walking at the same time.

To describe the dissociation between judgements about vagueness and typicality, it is common to associate concepts with *typicality functions*. Similarly to our treatment of acceptability functions, we will use typicality functions that map situations to a value between 0 and 1.⁹ The typicality value assigned to a situation may be different than the value assigned by the acceptability function. Summarizing the judgements in Examples 1–3 above, we make the following assumptions on the typicality and acceptability functions.

For the concept HAIR in Example 1, let S_1 be a situation with a focally red hair. Based on the observations above, we denote:

$$\text{ACC}_{\text{HAIR}}(S_1) \approx 1 \text{ and } \text{TYP}_{\text{HAIR}}(S_1) \ll 1.$$

For the concept SHAKE in Example 2, let S_2 be the situation in Fig. 2b. We denote:

$$\text{ACC}_{\text{SHAKE}}(S_2) \approx 1 \text{ and } \text{TYP}_{\text{SHAKE}}(S_2) \ll 1.$$

For the concepts WALK and WRITE in Example 3, let S_3 be a situation where Lucy is walking and writing simultaneously. We denote:

$$\text{ACC}_{\text{WALK}}(S_3) \approx 1 \text{ and } \text{TYP}_{\text{WALK}}(S_3) \ll 1, \text{ and similarly for WRITE.}$$

With this background on concepts, their vagueness and typicality effects, we can move on to the question of concept composition.

5 Guppy effects with gradable adjectives

In formal semantics it is commonly assumed that meanings of complex phrases are determined by a general principle, known as *compositionality* (Barker and Jacobson 2007; Janssen 1997; Werning et al. 2012). According to this principle, the meaning of a complex expression is determined by the meanings of its parts, and the way they combine with each other. For phrases like *a vegetarian student*, the simplest way to follow the compositionality principle is to use logical conjunction: a vegetarian student is someone who is independently categorized as both a student and a vegetarian. While in this example the conjunctive process works quite well,

⁹I here ignore the question of whether there is indeed a bounding value for typicality (Hampton 1998, 2007; Osherson and Smith 1981; Smith and Medin 1981; Smith and Osherson 1984). I also do not rule out the possibility that typicality and acceptability (at least at the lexical level) are determined by the same variable, as Hampton suggests. The data discussed in this paper are consistent with this possibility. The important point is the uncontroversial assumption that typicality rating may be different than acceptability.

other modification constructions require a more complicated semantic or pragmatic treatment. Osherson and Smith (1981) study the implications of applying the simple conjunctive treatment to compounds like *pet fish*, as in the following analysis.

- (5) An entity is categorized as a PET FISH if it is independently categorized as a PET and as a FISH.

According to O&S's intuitions, a guppy should be considered more typical for the category PET FISH than it is for any of the categories PET and FISH in isolation. This means that the typicality of an object for the category PET FISH cannot be easily determined using the object's typicality values for each of the two constituent categories independently. This kind of problem for understanding typicality was demonstrated in many experiments, starting in Smith and Osherson (1984). We collectively refer to such challenges about typicality as "guppy effects".¹⁰ Applying O&S's claims to adjective-noun compounds, Smith and Osherson (1984) showed that objects may be more typical for such constructions than they are for the adjective and the noun in isolation. For instance, in S&O's experiment, brown instances of apples received higher typicality ratings for the expression *brown apple* than for each of the words *brown* and *apple* in isolation.

Despite the relevance of O&S's puzzle for the study of complex concepts, it should be noted that O&S themselves did not consider guppy effects to be an immediate problem for a compositional rule as in (5). O&S's approach is based on a sharp distinction between typicality and vagueness (Osherson and Smith 1997). Thus, according to O&S's view, category membership can be determined using a compositional rule like (5), while typicality judgements are governed by other, possibly non-compositional, mechanisms. However, O&S's sharp dissociation of typicality from vagueness is not easy to maintain. As many works have shown, there are important relations between typicality and category membership. McCloskey and Glucksberg (1978) showed that speakers often disagree about category membership when exemplars have intermediate typicality values, but they are less likely to disagree about membership if typicality is very high or very low. Furthermore, Hampton (1988b) showed guppy effects for typicality rating tasks, as well as for membership rating tasks. For instance, Hampton studied membership rating for complex expressions like *sport which is also a game*. The results showed that the item *chess* was more often categorized as a *sport which is also a game*, than it was categorized as both a *sport* and a *game* independently. This kind of *overextension effect* challenges the conjunctive rule in (5), and it was demonstrated with many items in relative clauses. Chater et al. (1990) obtained similar results to Hampton's using a Y/N category membership task, without membership rating. Similar effects were shown with adjective modifiers, noun compounds, disjunction and negation (Hampton 1988a, 1997; Storms et al. 1996).

¹⁰Somewhat ironically, Storms et al. (1998) report a failure to experimentally attest a guppy effect with the *pet fish* example of O&S. However, this failure only concerns O&S's specific example, and the so-called guppy effect was attested in many experiments, including Storms et al.'s.

The overextension effects that were studied in the literature normally dealt with conjunctive expressions where there is not necessarily a conflict between the conjoined categories. For instance, *golf* is both a highly typical sport and a highly typical game, and in Hampton's experiments it was commonly judged to belong in both categories (Hampton 1988b). Thus, there is no necessary competition between the typicality that exemplars have with respect to the concepts SPORT and GAME.¹¹ A similar point holds for many of the items studied in the literature on overextension. By contrast, O&S were interested in concept composition where one concept is in a "negatively diagnostic" relation with another concept (Smith and Osherson 1984, p. 340). For instance, consider the inherent typicality conflict in the expression *red hair*. Hair instances that are typical of RED are atypical for the concept HAIR. Conversely, typical HAIR instances are of hues that are either not red at all or quite atypical for RED. Lee (2017) checked the effects of such typicality conflicts on truth-value judgements. His experiments contrast combinations like *red hair*, where there is an intuitive conflict in typicality (S&O's "negatively diagnostic" relation), and cases like *red car*, where there is no typicality conflict (S&O's "nondiagnostic" relation). Accordingly, in Lee's experiment 2, participants were asked to give their judgements on two images:

Image 1: an image of a woman whose hair is the same hue of Example 1 above (RGB value: 201,113,13—Lee, 2017: Fig. 5.4).

Image 2: an image of a car that is painted the same hue as in Image 1.

The majority (92%) of the participants in Lee's experiment accepted the categorization *red hair* for Image 1, but only 17% accepted the categorization *red car* for Image 2. This means that for at least 75% of the participants, a simple conjunctive analysis as in (5) would not work without additional assumptions. We can standardly assume that the likelihood that an object is categorized as *red* is affected by the linguistic and non-linguistic context.¹²

¹¹In fact, many typical sports are also typical games. For this reason, Hampton's (1997) study of constructions like *sports that are not games* is exceptional in that it (indirectly) tests potential typicality clashes, here between *sports* and *non-games*.

¹²For linguistic work, see Cresswell (1976), Kamp (1975), Kamp and Partee (1995), Klein (1980), Kennedy (2007). For experimental work, see Hansen et al. (2006), Kubat et al. (2009). Because of contextual effects, it is likely that some of the participants in Lee's experiment would accept the hair in Image 1 as *red* but reject the car in Image 2, even if the nouns would not be mentioned, relativizing redness to the visual instances of the concepts HAIR and CAR. Lee's work did not try to factor out possible effects of the visual stimuli, which are also central for concept composition (see Barsalou, 2017). However, Lee's effects in distinguishing gradable-concept categorization in complex expressions with "neutral" versus "biased" categorizes are much stronger than in previous work on categorization with simple color terms. Therefore, we can maintain the assumption that the interpretation of an adjective is affected both by its visual context and its linguistic context. Below I ignore effects of the visual context because they are less directly relevant to the compositionality problem in linguistics. Most importantly, effects coming from the non-linguistic context are not immediately testable when analyzing reciprocity and distributivity, because these concepts, unlike adjective concepts, are not easily studied in isolation: unlike category names such as RED, reciprocity and distributivity are bound to appear as sentence parts.

Here we concentrate on the effects of the noun concept on the interpretation of contrastive pairs like *red hair/car*. Such contrasts show a simple enough illustration of the problem for the conjunctive rule. Therefore, from a theoretical perspective, they are useful for analyzing the relations between membership and typicality with complex expressions. Like other gradable concepts, the concept RED imposes a natural ordering on entities. Here the ordering is naturally based on hue, and can be expressed by the comparative statement *x is redder than y*. Let us now consider the concept CAR in the phrase *red car*. The relation between the concepts RED and CAR is what S&O classify as a “nondiagnostic” relation: the typicality of CAR instances is likely to remain by and large unaffected by changes in hue. More formally, suppose that we are given two situations S_1 and S_2 with a car, where the only difference between S_1 and S_2 is in the car’s hue. We may reasonably assume that $\text{TYP}_{\text{CAR}}(S_1)$ is close to $\text{TYP}_{\text{CAR}}(S_2)$. Specifically, suppose that S_1 has a car painted focal red, and S_2 has the same car painted some other hue, quite distant from focal red. Both instances are expected to be equally typical for the concept CAR.

Unlike the concept CAR, the concept HAIR clearly has more typical and less typical colors. For instance, various shades of black and brown are more typical for human hair than, say, shocking pink. Among the hues between orange and focal red, some hues, at the margins of the concept RED, may be categorized as quite typical for HAIR. These are the hues that are most common for hair that is classified as RED HAIR. Let us informally refer to these hues as ‘ginger’. When we consider the complex expression *red hair*, we see a typicality conflict (S&O’s “negatively diagnostic” relation), which is due to the effect of hue on the typicality for HAIR. Starting from those hues that we called ‘ginger’, the redder the hue gets, the lower the typicality is for the concept HAIR. We classify this effect as *downward-monotonicity* of the typicality function for the concept HAIR, and say the function TYP_{HAIR} is downward-monotone relative to the order imposed by the gradable concept RED. More formally:

For any two instances x_1 and x_2 of HAIR, where x_1 and x_2 ’s hues are between ginger and focal red: if x_1 is redder than x_2 , then $\text{TYP}_{\text{HAIR}}(x_1) \leq \text{TYP}_{\text{HAIR}}(x_2)$

Note that this downward-monotonicity is only local: if we look at the hues that lie between the ginger hues and, say, the green hues, the HAIR typicality function is *upward-monotone*, since ginger hues are more typical as hair colors than green hues. Therefore, the ginger hues give a *local maximum* of the typicality for the concept HAIR among the hues that may reasonably be categorized as *red*.¹³

¹³The choice of green hues as the point where typicality starts to increase towards ginger is for presentation sake. Black, brown and blond may be globally more typical than ginger hues. Therefore, other hues besides ginger attain local maxima for HAIR typicality. However, between the black, brown and blond hues that are most typical for HAIR, we assume that typicality for HAIR must decrease somewhere before it raises again towards the ginger hues. Hence, for the sake of illustration, it is safe to assume that ginger attains a local maximum for hair typicality also in other dimensions of color other than the one illustrated here. This simplified presentation would not work for multi-dimensional concepts like PET or FISH, which introduce further complications when dealing with constructions like *pet fish* (Hampton 2007, p. 374). I believe that the current proposal can be

Above I have described the distinction between *red car* and *red hair* in terms of the behavior of the typicality functions for the concepts CAR and HAIR under changes of hue. Let us look at more concrete, hypothetical typicality functions that show this behavior, based on Lee's example. For purposes of illustration, we may suppose that green hues have typicality 0 for the concept RED, ginger hues (around RGB value 201,113,13, as in Lee's experiments) have typicality 0.6 for RED, and focal red has typicality 1 (Fig. 3 on top). In all our analyses, we concentrate on typicality for the *head concept* (HC): the concept that is being modified by the graded concept in the construction, e.g. the concepts HAIR and CAR in the constructions *red hair* and *red car*. For these two expressions, we look at situations that vary with respect to their typicality for the gradable concept RED. We encode situations according to their typicality for the concept RED. Specifically, we let the situations that we consider be described by numbers between 0 and 1, according to their typicality for RED. For instance, for RED HAIR we are interested in various situations where hair only differs in hue, and the hues range between green and focal red. Now, Fig. 3a approximates the typicality for the head concept HAIR as a function of the hue typicality for the gradable concept RED.¹⁴ As illustrated in this figure, ginger attains local maximal typicality for the concept HAIR. We look at this local maximum as an optimal "compromise" in typicality between the concepts RED and HAIR, referring to it as a *critical typicality* (CT) point for the complex concept RED HAIR.

Using the same kind of analysis, let us now consider the concept CAR. We assumed that typicality for CAR is constant in the hue. Thus, all hues obtain local maxima for CAR, as illustrated in Fig. 3b. Among these hues, we define the CT point for RED CAR as the *reddest* one, i.e. focal red. Thus, the CT point for RED CAR is much redder than ginger, which is the CT point for RED HAIR.¹⁵ Note that the difference between Fig. 3a and b is predicted from the difference between the typicality functions for the concepts HAIR and CAR, and the way they behave under changes in typicality for RED.¹⁶ Thus, our ability to see differences between expressions like *red hair* and *red car* is based on the information we have on their parts, in agreement with compositionality. However, unlike standard conjunctive rules, the information

(Footnote 13 continued)

subsumed by current models of distributional semantics (Baroni and Zamparelli 2010, Mitchell and Lapata 2010, McNally and Boleda, 2017). However, since *pet fish* examples concern problems with noun modification that are not immediately relevant for plurals, this task is left for further work.

¹⁴The graphs in Fig. 3, as well as in Figs. 4 and 5, are only for illustrative purposes. Using exact typicality values here is impossible, since the experiments by Lee, Poortman et al. and Poortman did not measure typicality directly, but only compared various situations as for their typicality. With respect to acceptability (as opposed to typicality), Figs. 7, 8 and 9 that follow later better approximate the results in those experiments.

¹⁵This idea is inspired by the following intuition in Krifka (1999): *Context-sensitive interpretation of RED can be given as: "When combined with a noun meaning N, it singles out those objects in N that appear closest to the color of blood for the human eye."*

¹⁶The process can be extended for complex phrases like *giant midget/midget giant* (Kamp and Pardee 1995), where both concepts are gradable (and unidimensional). However, the focus here is on those cases where one of the concepts is gradable and the other one is not, since in those cases the generalization to plural concepts is most direct.

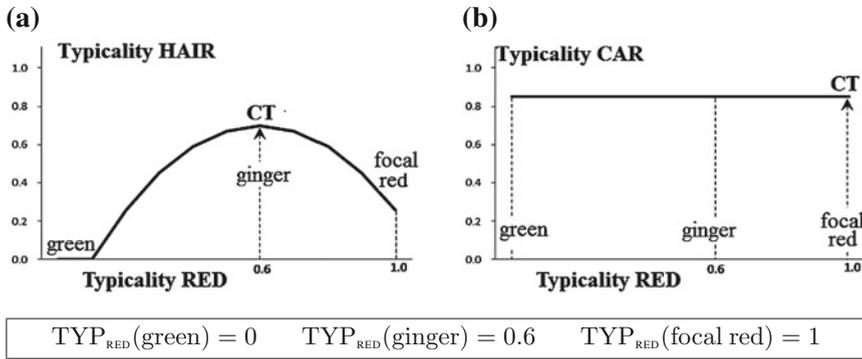


Fig. 3 Determining the CT point for RED HAIR and RED CAR: Typicality values of different hues for RED (see on top) are used to describe typicality for HAIR and CAR as a function of typicality for RED (a and b). The reddest local maximum point(s) of these functions is the critical typicality (CT) point (def. (7), example (6)). The CT point affects the determination of acceptability for the complex concept (Sect. 7)

we use is not just the extension of these parts, i.e. it is not the sets of items categorized by them. Rather, as the “meaning” of the parts we use the typicality functions derived from the lexical concepts.

With different CT points for the expressions *red hair* and *red car*, we are half way in the analysis of Lee’s results. As will be described more explicitly below, instances at the CT points are expected to be frequently judged as instances of the complex concept. Thus, since focal-red cars are at the CT point, they are likeliest to be classified as a *red car*. However, ginger cars are substantially below the CT point for CAR, hence they are unlikely to be classified as *red car*. By contrast, ginger hair instances are at the CT point for HAIR, hence are likely to be classified as *red hair*. A separate question, which will also be discussed below, is what happens with focal-red instances of HAIR, which are substantially *above* the CT points for HAIR. As we will see, different concepts behave differently at such points.

Before moving on to plural concepts, let us explicitly summarize our analysis of the *red hair* example.

- (6) Let LM-HAIR be the set of local typicality maxima for the concept HAIR. In formula:

$$LM-HAIR = \arg \max_x TYP_{HAIR}(x).$$

The *critical typicality* point(s) for RED HAIR is defined by:

$$\arg \max_{x \in LM-HAIR} TYP_{RED}(x).$$

In words: the critical typicality points for RED HAIR are those points that attain maximal typicality for RED among the points that attain maximal typicality for HAIR. Suppose that the local maxima in LM-HAIR are four points that represent the focal hues for brown, black, blond and ginger. Of these focal hues, ginger attains maximal typicality for RED, hence it is classified as the CT point for RED HAIR. A similar pro-

cedure yields a focal red as the critical point for RED CAR: supposing, as we did, that all hues are equally typical for CAR, we get the set of local maxima for CAR contain the whole spectrum. Of these hues, focal red is most typical for RED, hence classified as the CT point for RED CAR.

In more general terms, let us define the critical typicality point(s) for *gradable concepts* (e.g. RED) when they operate with *head concepts* like CAR or HAIR.

(7) **Critical Typicality (CT)**: *Let GC be a gradable concept and HC be a head concept. Let LM-HC be the set of local typicality maxima for HC, defined by:*

$$\text{LM-HC} = \arg \max_x \text{TYP}_{\text{HC}}(x).$$

*The set $\text{CT}_{\text{GC-HC}}$ of **critical typicality (CT) points** for the complex concept GC-HC is defined by:*

$$\arg \max_{x \in \text{LM-HC}} \text{TYP}_{\text{GC}}(x).$$

In words: the CT points for the complex concept GC-HC are those points that attain maximal typicality for the gradable concept GC among the points that attain maximal typicality for the head concept HC. This is precisely the same definition as in (6), but using more general notation.¹⁷ This definition will now help us analyze various guppy effects where the gradable concept is a *plurality operator*: distributivity or reciprocity.

6 Guppy Effects with Plurals

While guppy effects in concept composition have been widely studied in the literature on vagueness and adjectives, they have less often been addressed in relation to the interpretation of plurals. The main proposal of this paper is that the contextual mechanisms affecting meaning composition with vague adjectives are also operational with vague plural quantifiers. To see the parallelism, consider the following two sentences (Philip 2000; Winter 2001).

- (8) a. Mary, Lucy and Candy *are pinching each other*.
 b. Mary, Lucy and Candy *know each other*.

¹⁷James Hampton (p.c.) points out that definition (7) is not fully general. For instance, for GREEN HAIR, we would like the CT point to be around focal green, and not, say, ginger or focal black. However, since ginger and focal black are local maxima for HAIR typicality, definition (7) might classify them as CT points, despite their zero typicality for GREEN. To address this problem, we would need definition (7) to only consider local maxima of HC-typicality which are close enough to the maximal typicality points for GC. E.g., for GREEN HAIR, we would only consider hair instances at a certain distance d from focal green. Reasonably, all these points would be equally (a)typical for HAIR, hence classified as local maxima. Accordingly, focal green would be classified as the CT point for GREEN HAIR. By contrast, if we look at hair instances at distance d from focal *red*, we expect to see some increase in HAIR typicality at different points closer to the browns, blonds or gingers. This would only allow such points be the local maxima, in agreement with the analysis above. For sake of presentation, we avoid this complication of (7), which is unnecessary for the analysis of plural concepts.

As we may naturally expect, the concept PINCH shows preference for situations where an agent only pinches one patient at a time (this was experimentally verified by Kerem et al. and Poortman et al.). However, the reciprocity concept for *each other* prefers as many relations as possible (Dalrymple et al. 1998; Sabato and Winter 2012). In sentence (8a) this leads to a conflict between the preferences of the two concepts. By contrast, in sentence (8b) there is no conflict in preferences between the verb and the reciprocal expression: a person may know many people, without clear typicality preferences between different numbers of acquaintances. Intuitively, this contrast points to a possible guppy effect in sentence (8a) when compared to (8b). Indeed, Kerem et al. and Poortman et al. experimentally showed that there is a substantial difference in the interpretations of sentences like (8a) and (8b). Poortman et al. tested truth-value judgements on Dutch versions of these sentences (and similar ones) for I3 situations as in Fig. 1 above: three agents, each of them acting on another agent. While 88% of the participants accepted (8a) as true in this I3 situation, only 36% accepted (8b) in the same situation. Kerem et al. and Poortman et al. explain such differences as a guppy effect, using a principle that they call the *Maximal Typicality Hypothesis* (MTH). As an instance of the more general principle in (7), the MTH is analyzed as responsible for the guppy effect. For example, with the concept PINCH in (8a), the I3 situation is the most typical situation for the reciprocity concept that is consistent with the preference that one person does not pinch more than two people at the same time. If we try to add more relations to I3 in order to satisfy better the preferences of the reciprocal concept, we get atypical situations for the PINCH concept. As a result, I3 is the critical typicality point for the verb phrase *pinch each other* in sentence (8a). By contrast, in sentence (8b), with the concept KNOW, there is no substantial typicality difference between I3 and configurations containing more relations. Consequently, the preferences of the reciprocal concept are free to take over, and I3 is not a CT point for the expression *know each other* in (8b). Rather, in this case the CT point is the I6 situation, where every one of the three people knows every other person. Figure 4a, b summarizes the typicality considerations for sentences (8a) and (8b) with I3 and I6. For contrast, these figures also include the I2 situation, with only two pinching relations between the three people. In Kerem et al.'s and Poortman et al.'s experiments, I2 situations consistently showed very low acceptability for reciprocal sentences with all tested verbs.

In (9) below we apply principle (7) to the case of the complex concept for *pinch each other*, where PINCH is the head concept HC and RECIP is the gradable concept GC.

(9) Let LM-P be the set of local typicality maxima for the concept PINCH. In formula:

$$\text{LM-P} = \arg \max_x \text{TYP}_{\text{PINCH}}(x).$$

The CT point(s) for *pinch each other* is defined by:

$$\arg \max_{x \in \text{LM-P}} \text{TYP}_{\text{RECIP}}(x).$$

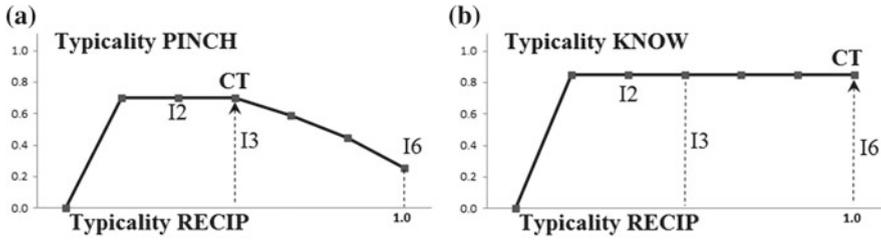


Fig. 4 Determining the CT point for *pinch each other* and *know each other*: Typicality values of the I2, I3, and I6 situations are used to describe typicality for PINCH and KNOW as a function of typicality for RECIP (a and b). The “most reciprocal” local maximum point(s) of these functions is the critical typicality (CT) point (def. (7), example (9)). The CT point affects the determination of acceptability for the complex concept (Sect. 7)

In words: the critical typicality points for *pinch each other* are the situations that attain maximal typicality for the gradable concept RECIP among the situations that attain maximal typicality for PINCH. As we saw, this entails that I3 is a CT point for *pinch each other*. This analysis is analogous to the analysis of the RED HAIR example. A similar procedure gives us I6 as the CT point for *know each other*, in analogy to the RED CAR example.

Poortman (2017) studies another guppy-like effect with plurals. Consider the following sentences.

- (10) a. Dan, Bill, John and George *are walking and writing*.
- b. Dan, Bill, John and George *are walking and singing*.

In these constructions, we view ‘bi-concepts’ like WALK&WRITE and WALK&SING as the head concepts. The gradable concept of “distributivity” DIST is not overtly present in the sentence, but is standardly introduced in the analysis, as discussed above. We analyze the distributivity concept DIST as a gradable vague concept similar to *many* or *enough*. The complex concept for the plural expression *are walking and writing* is denoted DIST WALK&WRITE (as opposed to the conjunctive concept WALK&WRITE for the uninflected conjunction). Similarly, the complex concept for the expression *are walking and singing* is denoted DIST WALK&SING.

Now let us consider three types of situations for sentences (10a–b), illustrated in Fig. 5:

- A *joint* situation, where each of the four people is doing both activities, i.e. walking and writing/singing.
- A *full-split* situation, where two people are performing one activity, and the other two are performing the other activity.
- A *partial-split* situation, where one of the persons is performing one activity, and another person is performing the other activity. The other two people are not engaged in any of the activities.

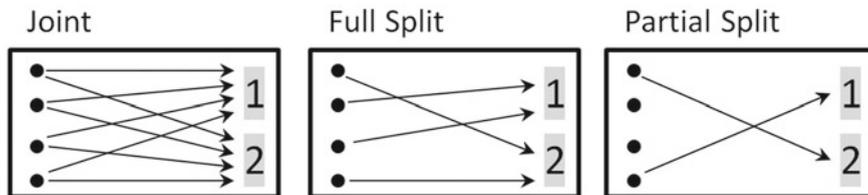


Fig. 5 Three situations for the interpretation of sentences (10a–b). The *dots* represent the agents Dan, Bill, John and George, and the *numbers 1* and *2* represent the activities done by each person (*walk, writelsing*, or both)

Partial-split situations are quite unacceptable for both sentences (10a) and (10b). Thus, when Dan walks, George writes and Bill and John are doing neither activity, sentence (10a) is quite odd, and similarly for (10b). By contrast, in full-split situations, Poortman’s experiments show a difference between Dutch sentences similar (10a) and (10b). While 81% of the participants accepted sentences like (10a) in a full-split situation, only 24% accepted (10b) (in Poortman’s experiment the subjects were definite descriptions, e.g. *the boys*). Once more, when accounting for this effect we rely on critical typicality points. In (11) below we apply principle (7) to the case of the expression *are walking and writing*. In this example, we denote WALK&WRITE for the head concept HC, where DIST, the distributivity concept, is the gradable concept GC.

- (11) Let LM-WW be the set of local typicality maxima for the concept WALK&WRITE. In formula:

$$\text{LM-WW} = \arg \max_x \text{TYP}_{\text{WALK\&WRITE}}(x).$$

The CT point(s) for the complex concept DIST WALK&WRITE is defined by:

$$\arg \max_{x \in \text{LM-WW}} \text{TYP}_{\text{DIST}}(x).$$

In words: the CT points for the expression *are walking and writing* are the situations that attain maximal typicality for the gradable concept DIST among the situations that attain maximal typicality for WALK&WRITE. Among the three situations we consider, the full-split and partial-split situations are more typical for WALK&WRITE than the joint situation. Between these two situations, the full-split situation is substantially more typical for the concept DIST, as it contains more people who are engaged in the relevant activities. This means that for sentence (10a), full-split is the CT point among the three situations. By contrast, for sentence (10a), the three situations are equally typical for the head concept WALK&SING. This is because the concepts WALK and SING, unlike WALK and WRITE, are not in any conflict (this kind of contrast is shown in another experiment by Poortman). In this case, for sentence (10b), the critical typicality point is the “joint situation”, which is the most typical for the distributivity concept alone. This analysis is summarized in Fig. 6.

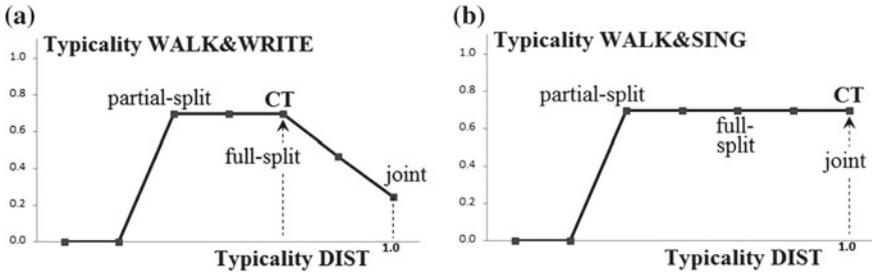


Fig. 6 Determining the CT point for *walking and writing* and *walking and singing*: Typicality values of different configurations (partial-split, full-split and joint) are used to describe typicality for WALK&WRITE and WALK&SING as a function of typicality for DIST (a and b). The local “most distributive” maximum point(s) of these functions is the critical typicality (CT) point (def. (7), example (11)). The CT point affects the determination of acceptability for the complex concept (Sect. 7)

7 The Effect of Critical Typicality on Acceptability

We have seen that for three different cases, typicality considerations allow us to specify points of *critical typicality*. Such CT points hit the optimal equilibrium between the (possibly conflicting) typicality preferences of concepts appearing together in a complex construction. We would like to use CT points in order to explain facts like the following:

- (12) a. Ginger-hued hair is more acceptable for *red hair* than car instances of the same hue are for *red car* (0.92 vs. 0.17 acceptability rates in Lee’s experiments).
- b. 13 situations are more acceptable for *pinch each other* than for *know each other* (0.88 vs. 0.36 acceptability rates in Poortman et al.’s experiments).
- c. Full-split situations are more acceptable for *walking and writing* than for *walking and singing* (0.81 vs. 0.24 acceptability rates in Poortman’s experiments).

The three experiments compared acceptability proportions when participants were asked to categorize instances of two complex expressions. From the perspective developed here, which follows Black (1937), a central goal of semantic theory is to account for such results on acceptability. It is important to note that *a priori*, it is not self-evident that acceptability proportions reflect typicality preferences for the complex concept. Typicality effects are only studied here insofar as they help to predict acceptability, as measured in truth-value judgement tasks. The typicality data that were tested in Lee’s, Poortman et al.’s and Poortman’s experiments concern the head concept, and not any complex concepts. As Fodor (1981) claimed, it becomes increasingly harder to study typicality information when it comes to com-

plex expressions.¹⁸ An example for the difficulty was illustrated in one of Kerem et al.'s 2009 experiments. In a forced choice experiment, only 26% of the participants preferred I3 to I6 as a better illustration for sentences of the form *A, B and C are pinching each other*. Thus, if we want to explain why I3 situations are more acceptable for *pinch each other* than for *know each other*, it is not clear how typicality effects might help us: when participants were asked about their preferences (as opposed to their acceptability judgements), many of them preferred I6 to I3 situations for *pinch each other*, and ignored the atypical status of such scenarios for the concept PINCH. This effect nicely demonstrates Fodor's point: as expressions get longer and more complicated, typicality judgements do not seem very informative.

Because of this point, we concentrate on acceptability judgements for complex expressions, rather than on typicality judgements. Typicality judgements are only collected for lexical expressions, which are then used for predicting the CT point(s). These CT points, in turn, are proposed to be an important factor that boosts acceptability judgements for complex expressions. The facts in (12) are all explained using the following generalization.

- (13) Let $GC-HC_1$ and $GC-HC_2$ be two complex concepts, with the same graded concept GC , but with different head concepts HC_1 and HC_2 . Let two situations, S_1 and S_2 be exemplars for HC_1 and HC_2 , respectively, where S_1 and S_2 are of equal typicality for the graded concept GC . Suppose that S_1 is a CT point for the complex concept $GC-HC_1$, but *below* the CT point for the complex concept $GC-HC_2$. Then the acceptability of S_1 for $GC-HC_1$ is substantially higher than the acceptability of S_2 for $GC-HC_2$.

For instance, since ginger hues are CT points for the concept HAIR, but not for the concept CAR, their acceptability for *red hair* is substantially higher than for *red car*. Similarly, since I3 situations are CT points for the concept PINCH, but not for the concept KNOW, their acceptability for *pinch each other* is substantially higher than for *know each other*. Further, since full-split situations are CT points for the concept WALK&WRITE, but not for WALK&SING, their acceptability for *are walking and writing* is substantially higher than for *are walking and singing*.

Note that the CT point is proposed to be a factor that boosts acceptability, but does not necessarily maximize it. Points above the CT point may have acceptability lower than or equal to the acceptability at the CT point, but they may also have higher acceptability. Here are two examples for this variability:

- In Lee's experiments, focal red hair was slightly less acceptable for *red hair* than ginger hair (0.87 vs. 0.92). By contrast, a bike made completely of wood was more acceptable for *wooden bike* than a more typical bike with some wood parts (0.96 vs. 0.71).

¹⁸Fodor used the complex phrase "American cities situated on the East Coast just a little south of Tennessee" to argue against the plausibility of prototype theory. According to Fodor, it is unlikely that such expressions have any prototype.

- In Poortman et al.'s experiments, the I6 situation was as acceptable for *pinch each other* as the I3 situation (0.88 in both cases). By contrast, the I6 situation was substantially less acceptable for *bite each other* than I3 (0.44 vs. 0.84).

Thus, above the CT point (I3, ginger hues, partially wooden bike), acceptability may increase, decrease, or remain the same. I propose that this happens because acceptability for a complex concept above the CT point is compositionally affected by the acceptability values for the parts. For instance, participants may fail to accept I6 situations as *biting* situations, simply because they reject the possibility that one person can be biting two people simultaneously. By contrast, a bike made completely of wood is still a bike, hence acceptability for *wooden bike* may continue to increase beyond the CT point.

Based on this account, we use CT points for specifying acceptability functions for complex concepts on the basis of the acceptability functions of their parts. Consider again our running example, the concept RED HAIR. Let us assume that x_0 is the least red hue that is still categorized as RED by some speakers.¹⁹ For instance, x_0 may be a slightly redder hue than focal green (where the latter has zero acceptability for RED). Based on the minimal acceptability point x_0 and the critical typicality point CT, we can describe the acceptability function for RED HAIR as consisting of three parts:

- (i) For every hue x between focal green and x_0 , the acceptability $ACC_{RED\ HAIR}(x)$ is constantly zero (as x is definitely not red).
- (ii) For every hue x between the CT point (ginger) and focal red: $ACC_{RED\ HAIR}(x)$ is compositionally determined by $ACC_{RED}(x)$ and $ACC_{HAIR}(x)$, e.g. by multiplication.²⁰ Intuitively, hair hues that are redder than the CT point(s) inherit their acceptability compositionally, from the acceptability for RED and for HAIR independently.
- (iii) For every hue x between x_0 and the CT point (ginger): $ACC_{RED\ HAIR}(x)$ grows monotonically—e.g. linearly.²¹ This penalizes acceptability of hues below the CT point, while allowing them to be non-zero.

This captures three facts about RED HAIR. Point (i) explains why instances of HAIR that are close enough to focal green, i.e. less red than x_0 , are fully unacceptable for RED HAIR, as they are for RED. Point (ii) explains how hair instances that are redder than ginger (the CT point) may have lower, equal, or even greater acceptability: the only restriction that this assumption puts on the acceptability function for the complex expression is that it is determined compositionally from the acceptability of its parts. Part (iii) of the acceptability function describes how points between x_0

¹⁹It is not guaranteed that such a point exists: acceptability for RED may asymptotically reach zero. The assumption of x_0 is only a convenience here.

²⁰The exact way to compose acceptability functions must be studied empirically, and is likely to be quite different from multiplication, which is only assumed here for illustration. The important assumption is that the acceptability of for the complex concepts grows monotonically in the acceptability of its parts.

²¹As with the case of multiplication above, linearity is only assumed here for illustration. The important assumption is that the function grows monotonically.

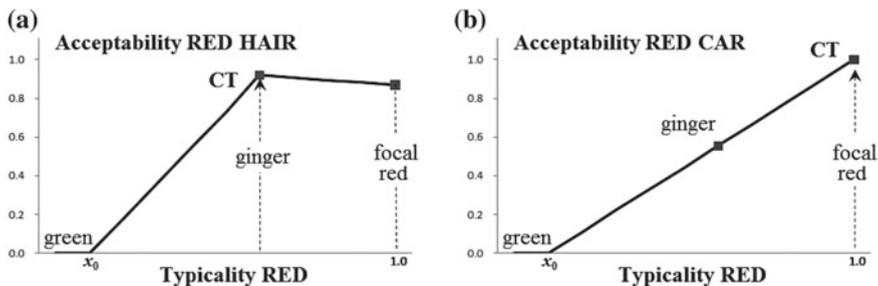


Fig. 7 Effects of CT point on acceptability of red hair and red car: The critical typicality (CT) point (def. (7)) is used for determining acceptability for complex concepts in two noun-adjective constructions (def. (14))

and ginger are only mildly acceptable for RED HAIR, while letting acceptability grow monotonically as hues move towards ginger.

In more general terms, we adopt the following rule, which defines how CT points interact with acceptability for complex concepts.

- (14) **CT-induced Acceptability:** Let GC be a gradable concept and let HC be a head concept. Suppose that in the interval $[0, 1]$, the CT point(s) for the concept GC-HC are precisely the segment/point $[x_1, x_2]$ where $x_1 \leq x_2$. Suppose further that x_0 is a maximal point s.t. $ACC_{GC}(x_0)$ attains a certain minimum, denoted *min*. Thus, for every $x \leq x_0$, $ACC_{GC}(x) = min$, and for every $x > x_0$: $ACC_{GC}(x) > min$. We define the acceptability function for the concept GC-HC as follows:

$$ACC_{GC-HC}(x) = \begin{cases} min & x \leq x_0 \\ \text{linear} & x_0 \leq x \leq x_1 \\ \text{constant} & x_1 \leq x \leq x_2 \\ x \cdot ACC_{HC}(x) & x \geq x_2 \end{cases}$$

The linear part is only assumed for concreteness, and may be replaced by any other function that grows monotonically in the acceptability for GC. The same remark holds for the last part, where multiplication may be replaced by any other two-place function that is upward-monotonic on both its arguments. The acceptability function $ACC_{RED HAIR}(x)$ of the concept RED HAIR is plotted in Fig. 7a, under the linearity assumption between x_0 and the CT point, and a slight decline in acceptability (in Lee’s experiment—from 92 to 87%) in the points higher than the CT point. In Fig. 7b, a similar graph describes the acceptability function $ACC_{RED CAR}(x)$ of the concept RED CAR. Here the CT point is the point for focal red, hence the acceptability function raises linearly from the point x_0 of minimal acceptability to that CT point of maximal acceptability. The difference between the graphs in Fig. 7a, b illustrates the effect of the CT point on the acceptability for complex concepts.²²

²²The linearity approximation in Fig. 7b grossly underplays the effect of the CT point in Lee’s experiment: in fact, only 17% of the participants accepted ginger hues for RED CAR.

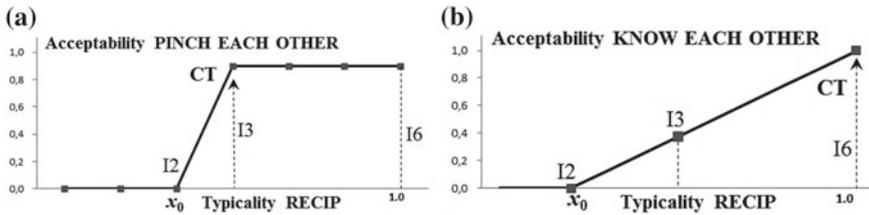


Fig. 8 Effects of CT point on acceptability of PINCH EACH OTHER and KNOW EACH OTHER: The critical typicality (CT) point (def. (7)) is used for determining acceptability for complex concepts in two reciprocal constructions (def. (14))

Let us consider another example for the application of CT-induced acceptability: the acceptability function for *pinch each other* in sentence (8a). We assume that $x_0=I2$ is a minimum point with zero acceptability for *each other*, and that the critical typicality point CT is I3.²³ As a result, we get the following acceptability function, with two parts in its definition (since I2 and I3 are discrete points).

$$ACC_{PINCH\ EACH\ OTHER}(x) = \begin{cases} 0 & x \text{ is a situation with 2 or less PINCH relations} \\ x \cdot ACC_{PINCH}(x) & x \text{ is a situation with between 3 and 6 PINCH relations} \end{cases}$$

Here the acceptability of any lower point than I3 (=CT point), with 2 or less relations, is zero (or otherwise very low, to the extent reciprocity tolerates such situations, see Beck and von Stechow 2007; Mari 2014). Higher points than the CT, between I3 and I6, inherit their acceptability from the acceptability of reciprocity (x) and the acceptability for the concept PINCH (which decreases with number of relations). In Poortman et al.’s experiments this leads to equal acceptability of I3 and I6 for *pinch each other*. The acceptability function for *pinch each other* is described in Fig. 8a.

When applying principle (14) to *know each other* in sentence (8b), we again use the assumption that $x_0=I2$ is a minimum point. However, as we saw, the CT point for *know each other* is I6, hence the acceptability function that (14) generates is different than for *pinch each other*. This function is given below.

$$ACC_{KNOW\ EACH\ OTHER}(x) = \begin{cases} 0 & x \text{ is a situation with 2 or less KNOW relations} \\ \mathbf{linear} & x \text{ is a situation with between 3 and 6 KNOW relations} \\ 1 & x = I6 \text{ (i.e. 6 KNOW relations)} \end{cases}$$

Here I2 and “smaller” situations are of low acceptability as with *pinch each other*. However, the CT point for *know each other* is I6. As a result, I6 is the only point where sentence (8b) attains maximal acceptability. In the points between I2 and I6, the acceptability function increases monotonically. This acceptability function for *know each other* is described in Fig. 8b.

²³The first assumption is only for the sake of illustration: actually 16% of the participants in Poortman et al.’s experiments accepted I2 for sentences like (8a). It is likely that situations like I1 and I0 (one and zero relations) would get even lower acceptability, hence the actual x_0 point is probably I1 or I0. The second assumption on typicality is based on the clear preference in Poortman et al.’s experiments for situations in which each agent only pinches one patient, rather than more than one.

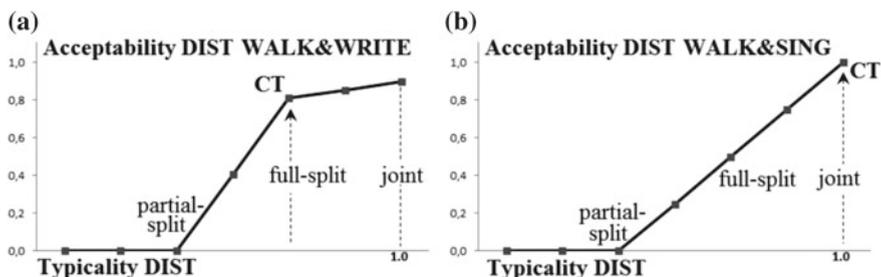


Fig. 9 Effects of CT point on acceptability of are walking and writing/singing: The critical typicality (CT) point (def. (7)) is used for determining acceptability for complex concepts in two conjunctive plural constructions (def. (14))

The analysis of the acceptability functions for plural predicates like *are walking and writing/singing* in (10a–b) is similar. For sentence (10a), we get the following acceptability function.

$$ACC_{DIST\ WALK\ \&\ WRITE}(x) = \begin{cases} 0 & x \text{ is any situation between partial to full-split} \\ & (=CT \text{ point}) \\ x \cdot ACC_{WALK\ \&\ WRITE}(x) & x \text{ is any situation between full-split and joint} \end{cases}$$

Here the CT point is the full-split situation, where two people are walking and two are writing. The acceptability of any lower point, with three or less active agents, is zero or very low (to the extent that distributivity tolerates such non-maximal situations). Higher points than the CT, between full-split and joint, inherit their acceptability from the acceptability of distributivity (x) and the acceptability for the head concept WALK&WRITE, which may stay constant or decrease a little with the size of the intersection, but (probably) not increase. The acceptability function for the complex concept DIST WALK&WRITE (for *are walking and writing*) is described in Fig. 9a.

For sentence (10a) the CT point is the joint situation, hence we get the following acceptability function.

$$ACC_{DIST\ WALK\ \&\ SING}(x) = \begin{cases} 0 & x \text{ is any situation between partial to full-split} \\ & (=CT \text{ point}) \\ \mathbf{linear} & x \text{ is any situation between full-split and joint} \\ x \cdot ACC_{WALK\ \&\ SING}(x) & x \text{ is the joint situation} \end{cases}$$

Here the partial-split situation and “smaller” situations are of low acceptability as with *walk and write*. However, the CT point is the joint situation, where plausibly, both distributivity and the WALK&SING concept attain maximal acceptability. In between these two points, the function increases monotonically. This function for the complex concept DIST WALK&SING is described in Fig. 9b.

8 Conclusion

This paper aimed to point out some general principles, which, if correct, should have desired implications for theories of concept composition in linguistics and psychology. It is well-accepted that truth-value judgements are systematically related to typicality, and that this relation is important for the compositional analysis of vagueness with nominal expressions. However, the general implications of this point for the analysis of compositionality are still understudied. Recent experimental work has shown that typicality effects influence the interpretation of plurals in a rather similar way to their relations with other vague concepts. Hopefully, the explorative work in this paper may help in improving existing theories of context and concepts, and the way they interact in the construction of meaning for complex expressions.

Acknowledgements I am grateful to James Hampton for his remarks on this paper. Thanks also to Philippe de Groote, Choonkyu Lee and Eva Poortman, for helpful discussions. Figure 2a, b were made by Ruth Noy Shapira. The work on this paper was supported by a VICI grant 277-80-002 of the Netherlands Organisation for Scientific Research (NWO).

References

- Alxatib, S., & Pelletier, F.J. (2011). The psychology of vagueness: Borderline cases and contradictions. *Mind & Language*, 26(3), 287–326.
- Barker, C., & Jacobson, P. (2007). Introduction: Direct compositionality. In C. Barker & P. Jacobson (Eds.), *Direct compositionality* (pp. 1–22). Oxford: Oxford University Press.
- Baroni, M., & Zamparelli, R. (2010). Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing* (pp. 1183–1193). Association for Computational Linguistics.
- Barsalou, L. W. (2017). Issues for psychologically plausible theories of conceptual combination. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 223–244). London: Springer.
- Beck, S., & von Stechow, A. (2007). Pluractional adverbials. *Journal of Semantics*, 24(3), 215–254.
- Black, M. (1937). Vagueness. an exercise in logical analysis. *Philosophy of Science*, 4(4), 427–455. Reprinted in Keefe & Smith (1996).
- Bonini, N., Osherson, D. N., Viale, R., & Williamson, T. (1999). On the psychology of vague predicates. *Mind & Language*, 14(4), 377–393.
- Brisson, C. (1998). Distributivity, maximality and floating quantifiers. PhD thesis, Rutgers University.
- Burnett, H. (2012). Vague determiner phrases and distributive predication. In M. Slavkovic & D. Lassiter (Eds.), *New directions in logic, language and computation* (pp. 175–194). Dordrecht: Springer.
- Chater, N., Lyon, K., & Myers, T. (1990). Why are conjunctive categories overextended? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 497–508.
- Cresswell, M. J. (1976). The semantics of degree. In B. Partee (Ed.), *Montague grammar*. New York: Academic Press.
- Dalrymple, M., Kanazawa, M., Kim, Y., Mchombo, S., & Peters, S. (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy*, 21, 159–210.

- Dowty, D. (1987). Collective predicates, distributive predicates and *all*. In *Proceedings of the Eastern States Conference on Linguistics, ESCOL3 (1986)* (pp. 97–115). Cascadilla Press.
- Egré, P., de Gardelle, V., & Ripley, D. (2013). Vagueness and order effects in color categorization. *Journal of Logic, Language and Information*, 22(4), 391–420.
- Fodor, J. A. (1981). The present status of the innateness controversy. In J. A. Fodor (Ed.), *Representations: Philosophical essays on the foundations of cognitive science* (pp. 257–316). MIT Press.
- Hampton, J. A. (1988a). Disjunction of natural concepts. *Memory & Cognition*, 16(6), 579–591.
- Hampton, J. A. (1988b). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(1), 12–32.
- Hampton, J. A. (1997). Conceptual combination: Conjunction and negation of natural concepts. *Memory & Cognition*, 25(6), 888–909.
- Hampton, J. A. (1998). Similarity-based categorization and fuzziness of natural categories. *Cognition*, 65(2), 137–165.
- Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science*, 31(3), 355–384.
- Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience*, 9(11), 1367–1368.
- Janssen, T. M. V. (1997). Compositionality. In J. Van Benthem & A. ter Meulen (Eds.), *Handbook of Logic and Language*. Amsterdam: Elsevier.
- Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, 57, 129–191.
- Kamp, J. A. W. (1975). Two theories about adjectives. In E. L. Keenan (Ed.), *Formal semantics of natural language*. Cambridge: Cambridge University Press.
- Keefe, R., & Smith, P. (Eds.). (1996). *Vagueness: A reader*. Cambridge, Massachusetts: MIT Press.
- Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30, 1–45.
- Kennedy, C. (2011). Ambiguity and vagueness: An overview. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 1, pp. 507–535). Berlin: De Gruyter.
- Kerem, N., Friedmann, N., & Winter, Y. (2009). Typicality effects and the logic of reciprocity. In E. Cormany, S. Ito, & D. Lutz (Eds.), *Proceedings of Semantics and Linguistic Theory, SALT19, eLanguage* (pp. 257–274).
- Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4, 1–45.
- Krifka, M. (1999). Compositionality. In R. A. Wilson, F. C. Keil, & A. Pierce (Eds.), *The MIT encyclopaedia of the cognitive sciences* (pp. 152–153). Cambridge, MA: MIT Press.
- Križ, M., & Chemla, E. (2015). Two methods to find truth-value gaps and their application to the projection problem of homogeneity. *Natural Language Semantics*, 23(3), 205–248.
- Kubat, R., Mirman, D., & Roy, D. (2009). Semantic context effects on color categorization. In N. Taatgen, & H. van Rijn (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (pp. 491–495). Cognitive Science Society, Inc.
- Langendoen, D. T. (1978). The logic of reciprocity. *Linguistic Inquiry*, 9, 177–197.
- Lasersohn, P. (1999). Pragmatic halos. *Language*, 75, 522–551.
- Lee, C. (2017). Color typicality knowledge and the interpretation of color adjectives. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 123–138). London: Springer.
- Mari, A. (2014). Each Other, asymmetry and reasonable futures. *Journal of Semantics*, 31, 209–261.
- McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, 6(4), 462–472.

- McNally, L., & Boleda, G. (2017). Conceptual versus referential affordance in concept composition. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 245–268). London: Springer.
- Mitchell, J., & Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science*, 34(8), 1388–1429.
- Nouwen, R., van Rooij, R., Sauerland, U., & Schmitz, H.-C. (Eds.). (2011). *Vagueness in communication*. Dordrecht: Springer.
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9, 35–58.
- Osherson, D. N., & Smith, E. E. (1997). On typicality and vagueness. *Cognition*, 64(2), 189–206.
- Pelletier, F. J. (2017). Compositionality and concepts: From a linguistics and philosophy of language perspective. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 31–94). London: Springer.
- Philip, W. (2000). Adult and child understanding of simple reciprocal sentences. *Language*, 76, 1–27.
- Poortman, E. (2017). Concept typicality and the interpretation of plural predicate conjunction. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 139–162). London: Springer.
- Poortman, E., Struiksma, M., Kerem, N., Friedmann, N., & Winter, Y. (2017). *When logic meets a prototype: Relational concepts and reciprocal reasoning* (to appear in *Glossa*).
- Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21–59). Cambridge: Cambridge University Press.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.
- Sabato, S., & Winter, Y. (2012). Relational domains and the interpretation of reciprocals. *Linguistics and Philosophy*, 35, 191–241.
- Sassoon, G. W. (2013). *Vagueness, gradability and typicality: The interpretation of adjectives and nouns*. Leiden: Brill.
- Schwarz, F. (2013). Maximality and definite plurals—Experimental evidence. In G. W. E. Chemla, & V. Homer (Eds.), *Proceedings of Sinn und Bedeutung*, 17 (pp. 509–526).
- Serchuk, P., Hargreaves, I., & Zach, R. (2011). Vagueness, logic and use: Four experimental studies on vagueness. *Mind & Language*, 26(5), 540–573.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, 8(4), 337–361.
- Storms, G., de Boeck, P., van Mechelen, I., & Ruts, W. (1996). The dominance effect in concept conjunctions: Generality and interaction aspects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1266–1280.
- Storms, G., de Boeck, P., van Mechelen, I., & Ruts, W. (1998). Not guppies, nor goldfish, but tumble dryers, Noriega, Jesse Jackson, panties, car crashes, bird books, and Stevie Wonder. *Memory & Cognition*, 26(1), 143–145.
- Werning, M., Hinzen, W., & Machery, E. (2012). *The Oxford handbook of compositionality*. Oxford: Oxford University Press.
- Winter, Y. (2001). Plural predication and the strongest meaning hypothesis. *Journal of Semantics*, 18, 333–365.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Complement Coercion as the Processing of Aspectual Verbs: Evidence from Self-paced Reading and fMRI

Yao-Ying Lai, Cheryl Lacadie, Todd Constable, Ashwini Deo
and Maria Mercedes Piñango

Abstract The so-called coercion verbs have been taken to select for an event as their complement, and to coerce an entity-denoting complement into an event as a resolution to the predictable type mismatch. This process is reported to manifest as additional processing cost that unpredictably has been associated with more than one cortical recruitment locus. Recent work has challenged the traditional view showing that the processing effect is observed only for aspectual verbs (e.g., *begin*) but not psychological verbs (e.g., *enjoy*) (Katsika et al. 2012), and that contra the traditional assumption aspectual verbs not only select for events but also for entity-denoting complements (Piñango and Deo 2015). Here, we test the hypothesis that aspectual verbs require their complement to be conceptualized as a structured individual. These verbs encode a set of functions that allow the construal of the structured individual as an axis along a dimension (e.g. *spatial, eventive*) afforded by the complement. The processing cost associated with the composition of the “coercion configuration” (animate subject + aspectual verb + entity-denoting complement) emerges from (A) exhaustive retrieval of the verbs’ lexical functions and (B) resolution of dimension ambiguity. Results from a self-paced reading and an fMRI experiment confirm that processing aspectual-verb sentences is more costly than psychological-verb counterparts, and that consistently with previous findings, comprehension is associated with both a Wernicke’s area and a left inferior frontal cortex activation. Crucially, this activation pattern tracks the necessary exhaustive lexical retrieval of the functions at the verb (Wernicke’s area) and the subsequent ambiguity resolution of the dimension at the complement (LIFG) required for the interpretation of the aspectual-verb utterance.

Y.-Y. Lai (✉) · M.M. Piñango

Department of Linguistics, Yale University, New Haven, USA
e-mail: yao-ying.lai@yale.edu

C. Lacadie · T. Constable

Yale Magnetic Resonance Research Center, New Haven, USA

A. Deo

Department of Linguistics, The Ohio State, Columbus, USA
e-mail: deo.13@osu.edu

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_8

1 Introduction

1.1 *The Complement Coercion Phenomenon*

It has been proposed that verbs like *begin* and *enjoy* carry a selectional restriction and must combine with an event-denoting complement (Pustejovsky 1995; Jackendoff 1997). Evidence in support of this restriction comes from the observation that even in sentences where the complement denotes an individual of the ordinary sort, an eventive interpretation is obtained. For instance, the sentence in (1), despite containing an entity-denoting complement, *the book*, can be interpreted as making reference to some event involving a book with John as its agent. In (2) and (3), the selectional restrictions of *begin* are satisfied when it combines with the gerundival complement *reading the book* or an event-denoting nominal *the fight*.

- (1) *John began/enjoyed the book.*
- (2) *John began/enjoyed reading the book.*
- (3) *John began/enjoyed the fight.*

This eventive interpretation associated with (1) has been interpreted in the literature as an instance of the broader phenomenon of **type coercion**.¹ The hypothesis is that there is a class of verbs that exclusively selects for event-denoting complements. This selectional restriction leads to a mismatch in the semantic representation when such verbs combine with complements denoting ordinary individuals. This mismatch is resolved by a semantic operation called *type-shifting* (Partee 1987; Partee and Rooth 1983) that coerces the semantic type of the entity-denoting complement into the appropriate event-denoting type (see Pyllkänen 2008 for a summary of descriptions of the hypothesis).

Experimental investigation of this phenomenon has revealed behavioral and neurological patterns that are taken to support this linguistic analysis. Psycholinguistic studies, using a variety of experimental paradigms, report that combining an entity-denoting complement with a coercion verb (*John began the book*) engenders more processing cost than combining it with a non-coercion verb (*John read/wrote the book*) during real-time comprehension (Baggio et al. 2010; Frisson and McElree 2008; Katsika et al. 2012; Lapata et al. 2003; McElree et al. 2001, 2006; Pickering et al. 2005, 2006; Scheepers et al. 2004, 2008; Traxler et al. 2002, 2005). On the neurolinguistic side, it has been found that the complement coercion effect recruits three distinct cortical regions: Wernicke's area (Piñango et al. 2001), ventro-medial

¹The term *coercion* is widely used to describe diverse phenomena in which it appears that interpretations are derived despite apparent semantic incongruity or mismatch between the combining expressions. The phenomena often considered under the coercion umbrella include nominal coercion with mass-count nominals (e.g. Pustejovsky 1995; Michaelis 2005; Wiese and Maling 2005), aspectual coercion with grammatical aspect (De Swart 1998) and aktionsart (e.g. Piñango et al. 1999; Todorova et al. 2000, Piñango and Zurif 2001; Piñango et al. 2006, Deo and Piñango 2011), and our present focus, complement coercion with complements of certain verbs.

prefrontal cortex (vmPFC) (Pylkkänen and McElree 2007; Pylkkänen et al. 2009), and BA45 (Husband et al. 2011). The additional cost has been interpreted as manifesting the type-shifting operation or the mechanism of building an eventive representation from an entity-denoting expression (Frisson and McElree 2008: 8; McElree et al. 2001: 22; Pickering et al. 2005: 9). The observed patterns in particular brain regions have been interpreted as evidence that these areas implement the operations of type-shifting and coercion.

1.2 Challenges to the Type-Shifting Hypothesis

The type-shifting hypothesis is, at its core, a hypothesis about the lexical semantics of so-called “coercion verbs” and the conceptual entities that composition involving such verbs is sensitive to. We propose that this hypothesis as formulated faces at least three challenges which have direct implications for how the experimental results are interpreted.

The *first challenge* has to do with the underlying linguistic pattern. For at least aspectual verbs, the most representative subset of coercion verbs, the eventive interpretation is not obligatory in the presence of an entity-denoting complement. In their simple transitive uses, the context in which these verbs have been tested, aspectual verbs do not exclusively select for eventive complements and agentive subject-referents. Sentences (4)–(6), found in the Corpus of Contemporary American English (COCA), exemplify such cases. For example, there is no construal of (4) in which *the book* is coerced into an event of which *the new autobiographical memoir* is an agentive participant. The same is observed in (5); the sentence does not give rise to an eventive interpretation though the “coercion verb” is followed by an entity-denoting complement (*the genealogy of the kings of England*).²

- (4) *Although this is mostly a collection of previously published essays, it is notable because of the new autobiographical memoir that begins the book.*
- (5) *This image begins the genealogy of the kings of England and flows into materials specifically written for St. Albans.*
- (6) *This column continues the review of the new PSA Club Services Website.*

²A reviewer points out that the meaning of “*begin*” in these cases is similar to “open up” or “be the beginning of X,” which would not be possible in (1) *John began the book*, thus leading to the possibility of two distinct uses of “*begin*.” On this point we observe this “open up” interpretation of “*begin*” can also apply to (1), giving rise to a metonymic reading such as “The story of John began the book.” This is what we call the constitutive interpretation of sentences involving aspectual verbs with an animate subject. Under our proposed Structured Individual analysis, it is one of the plausible interpretations of the sentence (in addition to the more salient eventive interpretation paraphraseable as “John began reading the book”). So, under this view both readings are accounted for. Details are discussed in Sect. 1.3.

This observation has been elaborated on in Piñango and Deo (2015) who observe that “not only is there no eventive complement (syntactically explicit or coerced) in these examples, but the sentences themselves are aspectually stative, and it is more accurate to say that they report configurational relations between individuals rather than causal relations between events (*p.* 10).” These distributional and interpretational patterns for the subset of aspectual verbs cast doubt on the validity of the empirical generalization that underpins the type-shifting analysis.

The *second challenge* comes from distributional differences in the complements of coercion verbs. Reporting the results of a corpus-based study, Utt et al. (2013) found that aspectual verbs co-occur significantly more with event-denoting nominals (such as *fight* in (3)) than psychological verbs. They call such expressions event-nouns and define them as nouns denoting actions, cognitive processes, or biological processes. These findings, while not *incompatible* with an event-selecting lexical semantics for coercion verbs, are rather surprising. There should be no reason why psychological verbs, if they lexically select for event-denoting complements, occur far less frequently with such complements than aspectual verbs. Further, this differential frequency distribution suggests that if the type-shifting hypothesis is correct, then the nominal complements of psychological verbs must undergo coercion far more frequently than the nominal complements of aspectual verbs. This asymmetry between the two classes is not naturally reconcilable with the uniform event-selecting lexical semantics for aspectual and psychological verbs that the type-shifting hypothesis relies on.

Finally, the *third challenge* comes from real-time implementation of coercion. Specifically, Katsika et al. (2012) observe that the class of coercion verbs investigated in previous studies on complement coercion was not only semantically heterogeneous, including at least two distinctly separable subclasses—aspectual verbs (e.g. *begin, finish, start*) and psychological verbs (e.g. *enjoy, prefer, endure*),³ they also observed that when these two subclasses were separately studied, they exhibited distinct processing profiles. The coercion *effect* (increased cost sometime after the complement head) was observable with the aspectual verb set and not with the psychological verb set.⁴

³The full class of coercion verbs investigated (in conflated fashion) also includes verbs like *master* and *attempt*. See Katsika et al. (2012).

⁴Here we note that the psychological verb set examined in coercion studies only included a small set of possible psychological verbs guided by the V + V-ing paraphrase diagnostic: John enjoys the book = John enjoys reading the book. Crucially, even though all psychological verbs pass the diagnostic, the diagnostic brings out the eventive interpretation more saliently with some psychological verbs (e.g. *enjoy, endure*) than others (e.g. *love, like, hate*).

These three challenges, taken collectively, suggest that the type-shifting hypothesis, and the lexical semantics for coercion verbs that it presupposes, is at best, problematic. In terms of grammatical possibilities, frequency distribution, as well as processing behavior, coercion verbs do not appear to form a unified class.⁵

In what follows, we present a hypothesis which on the basis of a new analysis for aspectual verbs (Piñango and Deo 2015) captures the pattern of real-time comprehension of aspectual verbs observed so far and makes further testable predictions; predictions which we then proceed to test from processing and neurological perspectives.

1.3 *The Structured Individual (SI) Analysis*

Piñango and Deo (2015) propose that aspectual verbs lexically select for complements that are **structured individuals**, rather than events. This proposal is motivated by the goal of offering a unified analysis for both “coercion configuration” sentences like *John began the book* and sentences like *The chapter on global warming began the book* (similar examples were given (4)–(6), not part of the traditional coercion set). The intuition underpinning a structured individual is an ordinary entity that maps onto a one-dimensional directed path structure (one-dimensional DPS) along a range of dimensions. A one-dimensional directed path structure is defined, following Krifka (1998), as a totally ordered structure whose adjacent (non-overlapping) parts bear the precedence relation along some dimension (temporal, spatial, eventive, etc.). For instance, one may construe an

⁵Two anonymous reviewers argue that these challenges do not necessarily falsify the type-shifting analysis, as the analysis can still be maintained by restricting the coercion phenomenon to the set of aspectual verbs and excluding psychological verbs and the remaining heterogenous set of verbs. One reviewer specifically suggests that the type-shifting account can restrict its claims to the combination consisting of an animate subject, an aspectual verb, and an entity-denoting object. If thus restricted, the processing cost can be accounted for through the following two steps: Step 1: type-shifting of the entity-denoting object, creating multiple possible eventive interpretations; and Step 2: event ambiguity resolution involving the selection of the proper event in context. We note two problems with this general option: First, the fundamental assumption on which the type-shifting analysis rests, which is that coercion verbs select for eventive arguments, is not valid, as demonstrated by examples (4–6). Without this assumption there is no motivation for type shifting (since the claim is that this mechanism is triggered by the presence of a mismatch). Hence, the processing cost induced by aspectual verbs (see Exp.1) remains unaccounted for as it cannot be attributed to the type-shifting operation (Step 1). As for Step 2, previous studies have already shown that the processing cost cannot be attributed to the selection of the right kind of event among plausible ones. Traxler et al. (2005) report that introducing a probable event prior to the target coercion sentence did not attenuate the cost. Also, Frisson and McElree (2008) indicate that coercion expressions with several alternative interpretations and no dominant one were not more taxing than expressions with a strong dominant interpretation. In their words, “one cannot reduce the cost of coercion to the more general effects of ambiguity or competition between different interpretations” and that the effects are “not likely to reflect costs associated with deriving an activity for the event sense.” (Frisson and McElree 2008: 8).

entity like a bridge, which is a three-dimensional spatial entity, as mapping onto a line (a one-dimensional DPS) along its most *salient* dimension, say spatial (Jackendoff 1992; Verkuyl and Zwarts 1992). Similarly, the text of a book may be construed as a one-dimensional path structure along the informational dimension, taking its chapters as the adjacent parts. More examples from the COCA are given below further illustrating the intuition.

- (7) *A moving train finishes the display.* (Spatial part)
- (8) *This January 1 begins the dawn of a new age of attainable resolution.* (Temporal part)
- (9) *His death begins the Revel.* (Eventive part)
- (10) *The only adequate or appropriate response to this reality seems to be the expression that both begins and ends the novel.* (Informational part)

Following Gawron (2009) and Deo et al. (2013), Piñango and Deo (2015) call such one dimensional DPSs in any ontological domain an “**axis**.” In defining the notion of a structured individual, they make reference to axes onto which individuals are mapped. This simply means that the individual is in a homomorphic relation to an (ontologically, conventionally, or pragmatically) given one-dimensional DPS. The predicate **axis** is taken to be the set of all entities (temporal, spatial, material, or abstract) that are one-dimensional DPSs. An axis is an element from this set.

Structured individuals are defined as entities that are relatable to such axes via homomorphic functions. The formal definition is given in (11). According to (11), an individual x of any type τ is taken to be a structured individual relative to a function f of any type (τ, σ) iff $f(x)$ is an axis and f is a homomorphism from the part structure of x to the axis $f(x)$.

- $$(11) \quad \forall x_{\tau} [\text{struct-ind}_{f(\tau, \sigma)}(x) \leftrightarrow [\text{axis}(f(x)) \wedge \forall x', x'' \leq x [x' \leq x'' \leftrightarrow f(x') \leq f(x'')]]]$$
- (Piñango and Deo 2015)

Aspectual verbs are analyzed as having both a presuppositional and truth-conditional component. They carry a lexical presupposition that requires their complement denotation to be a structured individual. Truth-conditionally, they require the subject to map onto some privileged small subpart of the axis determined by the complement denotation. A sentence with an aspectual verb is true iff the subject denotation is construed as a specific (e.g. initial, medial, final) subpart of the axis that the complement denotation is mapped to.

The sample lexical entry for *begin* in (12) (P & D 2015), illustrates this general schema assumed for aspectual verbs. The function (f_c) that maps the complement denotation to the axis comes from a (lexically encoded) set of functions each associated with a dimension. The idea encapsulated in the definition is that aspectual verbs are sensitive to whether their complement denotation can be construed in context as a structured individual, i.e. whether it is possible given contextual information to map it to an axis along some dimension. The contextually

accessed function f_c reflects this contextual reliance. Once the presupposition that the complement denotes a structured individual is met, aspectual verbs require that there be some function f' ($\exists f$) such that it maps the subject denotation to a small initial part ($<_{\text{small-init}}$) of the axis given by $f_c(x)$.⁶

- (12) a. $[[\text{begin}]] = \lambda x_\tau \lambda y_\sigma: \mathbf{struct-ind}_{f_c}(x). \exists f [f(y) <_{\text{small-init}} f_c(x)]$
 b. $\text{Begin}(x)(y)$ is defined iff x is a structured individual with respect to the contextually determined function f_c . If defined, $\text{begin}(x)(y)$ is true iff there is some function f (possibly identical to f_c) such that $f(y)$ is a small initial subpart of the axis $f_c(x)$.
 (Piñango and Deo 2015)

With this generalized lexical meaning, aspectual verbs are able to combine with arguments of different semantic types to yield a range of possible interpretations.

Consider the example in (7):

- (7) *A moving train finishes the display.* (Spatial part)

In (7), the complement (*the display*) is construable as a structured individual along the spatial dimension, mapping via the spatial trace function σ to its spatial extent—yielding an axis along the spatial dimension—the spatial extent of the display. The sentence is true iff the spatial extent of the entity denoted by *a moving train* is a small final subpart of the axis that the display maps onto along the spatial dimension.

The selection of the functions that map the complement denotation and the subject denotation to the axis is constrained by both the properties of the complement and context. The lexical meaning of the complement must encode information about the possible dimensions that are relevant to understanding the concept denoted by the complement. For instance, an expression like *magazine* must contain information about its spatial as well as informational structure; expressions like *bridge* and *river* must contain the information that they have a salient spatial dimension. This information will be mined during composition with aspectual verbs

⁶Notice that it is not necessary that the subject denotation and the complement denotation be mapped to the axis by the same function. The complement denotation maps to the axis by the contextually given function f_c while the function that maps the subject denotation to the axis may be distinct from it (hence the existential quantification over the function that associates the subject denotation to the axis). This separation between the two functions is motivated by examples like “*A veritable lineup of our planetary neighbors begins the month of March,*” where the subject denotation is best understood as an event while the complement denotes a temporal interval. In this case, the temporal dimension from the complement (obtained by the identity function) is mapped to the axis, while the function that maps the subject denotation to the axis is the runtime function τ , that maps events to their runtime. The coercion configuration also maps the complement denotation and the subject denotation to the eventuality axis via non-identical functions, as we describe later in this section. See Piñango and Deo (2015) for the detailed implementation.

(and probably in other cases as well), for determining the dimension along which the complement is to be construed as a structured individual in a given context.

The structured individual analysis treats the meanings of sentences containing aspectual verbs to be underdetermined, with full determination being dependent on which dimension is chosen for interpretation in a given context. Take “*The tedious preface begins the novel*” for an example. Here the complement can be conceptualized as a body of informational content—an axis along the informational dimension. *The tedious preface* maps to its informational content and is asserted to be a small initial subpart of the informational content of the novel. In the informational dimension, the physical form (e.g. size, number of pages) of the novel is irrelevant. It does not matter which page the content of preface appears on; the preface could be printed on page 5, preceded by the title page, the copyright page, or the table of contents. Since the truth of the sentence is determined by the informational extent of the two arguments, it is possible that the sentence is false on the spatial interpretation, in which the preface is printed on page 5. On the other hand, the novel can be conceptualized as a spatial entity—an axis along the spatial dimension, in which the novel denotes a structured individual consisting of physical pages of a particular size, with a pagination order. In the spatial dimension, the sentence is true iff the preface is printed on the first page of the novel. A printer might say something like “*It is weird that, instead of the half-title and the copyright page, it is the preface that begins the novel.*” In this context, no reference is made to the informational content of the preface.

Given that entity-denoting expressions may map onto axes in multiple dimensions, comprehenders must rely on the context to determine the exact function that determines the dimension along which the axis is construed. In the context of formatting or printing the novel, the spatial dimension will be chosen; in the context of reading the novel, the informational dimension is more relevant, and therefore the more salient dimension.

The “coercion” uses of aspectual verbs are naturally accommodated within this general analysis. This configuration is characterizable as one in which an aspectual verb combines with an animate agentive argument and an entity-denoting complement with the resulting reading that the animate subject referent is the agent of some implicit dynamic eventuality. These cases are analyzed as involving functions that map entities to the events that they are participants of—inverse thematic functions. While thematic roles (e.g. *agent*, *patient*, *theme*) map events to their participants as the actor or undergoer, inverse thematic functions are defined as functions that “map pairs of individuals and times to the smallest event that the individual bears a participant role to at that time in a given context” (Piñango and Deo 2015). So in an intuitive way, inverse thematic functions are ways of accessing events via the individuals that participate in them, rather than accessing individuals via the participant role that they bear in an event. f_{ag_i} maps an individual to the smallest event that they are the agentive participant of at the reference interval i in a

given context. f_{th_i} maps an individual to the smallest event that they are the patient/theme of at i in a given context.

The salient reading for a sentence like “*Jane Austen began the anthology*” is one that involves an eventive dimension (*Jane Austen began reading/writing/etc... the anthology*). In this case, the axis is construed as some event of which Jane Austen and the anthology are participants. The function f_c is taken to be f_{th_i} , which maps the complement, the anthology (as a structured individual), to the smallest event of which it is the theme at the reference time. The subject, Jane Austen, is mapped to the smallest event of which it is the agent by the inverse thematic function f_{ag_i} . The sentence comes out true if the smallest event of which *Jane Austen* is an agentive participant at a reference interval i in a given context is a small initial part of the event of which the book is a theme participant at i . This is the **agentive** reading of the sentence.

However, the sentence also has another, **constitutive**, reading where the individuals, *Jane Austen*, and *the anthology*, get mapped to an informational axis, such that the informational content corresponding to Jane Austen is understood to be a small initial subpart of the informational content corresponding to the anthology. In order to achieve this interpretation, metonymy has to be applied onto the subject, yielding an interpretation that may be paraphrased as [*the work of Jane Austen*] *began* the anthology.

Although one reading might appear to be more salient than the other, the agentive and constitutive readings are both possible in principle and are determined based on context and plausibility. Thus the indeterminacy of interpretation (agentive and constitutive) also extends to aspectual verb sentences in the coercion configuration.

The underspecification of meanings for aspectual verbs in the proposed analysis may give rise to a concern that the Structured Individual analysis may over-generalize. Indeed, a reviewer notes that sentences like “*The jogger began the bridge*” sound awkward but are, in fact, predicted to be acceptable by our analysis, given the spatial structure of the bridge. We agree that such sentences are predicted to be acceptable and note that, given the right context, such readings do become available. Consider, for example, the web-attested sentences (13–14).

- (13) *Fred Sebolt began to get rocks for the Wagon Bridge at Liberty Falls in the last week in August, began the bridge in September 1881.* (→ began building the bridge)
- (14) *So with 25 miles left my teammate and I attached and began the bridge in a headwind.* (→ began cycling across the bridge)

Here, we note that the Structured Individual analysis, in fact, *constrains* the properties of the eventuality that is construed as the relevant axis in the coercion configuration. This is because the structured individual presupposition requires that the contextually determined f_s be a homomorphism from the part structure of the complement denotation to the part structure of the axis. The axial eventuality must therefore be one in which the part structure of the complement denotation is

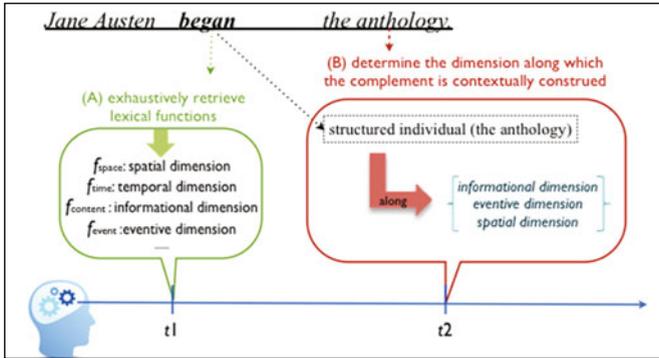
incrementally related to the course of the event—that is, the complement of *begin*, *finish*, etc. must be interpreted as an incremental theme argument of the implicitly construed eventuality. This is indeed the case with coercion sentences like *John began the book*—which cannot make reference to an event in which the book is not construed incrementally (such as an event of John playing with or seeing the book). This incrementality constraint on the interpretation of the eventuality associated with aspectual verbs in the coercion configuration falls out as a natural consequence of the lexical meanings for aspectual verbs assumed in the Structured Individual analysis—but must be stipulated in a type-shifting semantics for these verbs.

To summarize, the Structured Individual analysis not only grounds the lexical meanings proposed (i.e., dimension functions) on independently motivated conceptual properties, it makes no assumptions about selectional restrictions with respect to events. Therefore, under this analysis, the comprehension of aspectual verbs requires neither a type mismatch nor the implementation of a type-shifting operation. Instead, under this analysis, aspectual verbs combine with their complements and subjects just like other transitive verbs, but their full interpretation requires the contextual resolution of the specific dimension along which the complement can be construed as a structured individual. We call this process the resolution of dimensional ambiguity.

The analysis leads to a hypothesis about the processing of aspectual verbs which we term the **Structured Individual (SI) hypothesis**. On this hypothesis, the observed psycho- and neuro-linguistic reflexes of complement coercion are taken to reflect not type-shifting operations but the retrieval of the potential dimension-functions and the ultimate dimensional resolution. This allows us to not only maintain a uniform semantics for aspectual verbs across its uses (viz. agentive and constitutive readings), but also captures the observation from Katsika et al. (2012) that aspectual verbs, but not psychological verbs, engender additional cost.

Under the Structured Individual hypothesis, the processing of aspectual verbs is implemented in real-time comprehension as follows:

- (A) When readers encounter an aspectual verb, they retrieve the verb containing a number of (lexically encoded) dimension functions. We call this process the **exhaustive activation of lexically encoded functions**. These functions map between the domain of individuals (denoted by the subject) and subparts of the axis construed from one of the dimensions associated with the complement denotation.
- (B) In order to get a determinate interpretation for the composition of the aspectual verb and the complement, readers must determine the dimension along which the complement denotation should be construed as a structured individual. This means in turn, that the parser must choose a particular function from among those encoded in the verb. Because each complement denotation provides multiple dimensions, the parser is faced with dimensional ambiguity; an ambiguity that must be resolved for the sentence to be interpreted at all. Once the function is chosen (out of the possible ones offered by the interaction of complement and subject's denotations), the complement can be construed



Schema 1 Processing of aspectual verbs

as a structured individual along the dimension associated with that function. This is an instance of **ambiguity resolution**. These two processes are described in Schema 1.⁷⁸

According to this analysis then, resolving *any* sentence containing an aspectual verb requires the above two processes; processes that have been previously and independently invoked to capture semantic composition effects: (A) exhaustive activation of the verb’s lexical functions (e.g. Shapiro et al. 1989⁹), and (B) resolution of ambiguity created by immediate composition demands, i.e. dimension extraction from the complement (e.g. Frazier and Rayner 1990¹⁰).

⁷As Schema 1 shows, the ambiguity resolution involves deciding between the dimensions along which the structured individual associated with the complement denotation will be construed. Crucially, one of these dimensions can be eventive, leading to the standard coercion reading. As Frisson and McElree (2008) have shown, the kind of ambiguity between, say, *reading the anthology* versus *writing the anthology*, is not expected to contribute to cost.

⁸The lexical functions and the ontological dimensions listed here are not exhaustive. Piñango and Deo (2015) indicate that the dimensions can be “more abstract than that of space, time, or pieces of text.” Consider their examples (web-attested):

- (i) *Black starts and ends the visible spectrum.*
- (ii) *... the 6 neighborhoods based on social status (on a scale of 1 to 6): 1 is the poorest and the most dangerous of the neighborhood, 4 begins the middle class...*

Sentence (i) makes reference to “an ordering of electromagnetic radiation corresponding to the visible spectrum by wavelength” and (ii), to social status.

⁹This study shows that verbs with more possible argument structure arrangements (dative verbs) increased response time to the secondary lexical decision task than those with fewer possible arrangements (transitive verbs), regardless of the context. The authors therefore suggest that all the argument structure information of a verb must be momentarily activated during real-time processing.

¹⁰They found that delaying the disambiguation information for sentences that contain a word with multiple meanings increased processing cost.

We propose here that the “coercion” cost previously observed arises from the processing of aspectual verbs, the best represented verb class tested among the verbs tested in the past. In addition, we hypothesize that the seemingly incongruent neurological patterns previously associated with complement coercion, are in fact consistent with both step (A) and step (B).

In what follows, we present the results of two experimental studies that examine the psychological and neurological viability of the Structured Individual hypothesis (and ultimately the Structured Individual analysis). To this end, we carry out two experiments using self-paced reading (Exp. 1) and fMRI (Exp. 2) respectively, along with a pretest questionnaire. Three verb types are considered: aspectual verbs (AspectualV), psychological verbs of the *enjoy*-“type” (EnjoyingV) (these include only those psychological verbs that were tested as coercion verbs in previous work (see Katsika et al. 2012 for details)) and psychological verbs of the *love*-“type” (LovingV), which have been previously claimed to involve no coercion (Pustejovsky 1995) and thereby serve as controls. The three conditions all contain sentences with an animate subject and an entity-denoting complement.

Based on the Structured Individual hypothesis, we predict that aspectual verbs will induce longer reading times than either type of psychological verbs in the self-paced reading because the former involve the resolution of dimension ambiguity. With respect to neural correlates, we expect aspectual verbs to recruit additional cortical areas at two positions corresponding to the two processes mentioned above: (a) when the subject combines with an aspectual verb, where the parser exhausts the verb’s lexical functions, and (b) when the aspectual verb combines with the complement, where the parser must mine the complement denotation to determine the possible dimensions along which it is construable as a structured individual (Schema 1).

2 Pretest: Norming Questionnaire

To ensure equal acceptability of the manipulated conditions, we employed a rating questionnaire to test the stimuli.

2.1 Method

2.1.1 Participants

Forty native speakers of American English took the questionnaire, all between the ages of 18–30 and without reading disabilities. The data of three participants were discarded because their responses were either undifferentiated or inconsistent.

Table 1 Conditions and sample sentences

Condition	Example sentences
	<i>Lady Gaga</i>
AspectualV	<i>started</i>
EnjoyingV	<i>preferred</i>
LovingV	<i>loved</i>
	<i>this CD of American pop hits.</i>

2.1.2 Materials

We created 50 triplets; each containing the three manipulated conditions as shown in Table 1. Aspectual verbs (AspectualV) were contrasted with psychological verbs of the *enjoy* type (EnjoyingV). Note that the two are collapsed as “coercion verbs” under the type-shifting account. We further introduced psychological verbs of the *love*-type as the control condition (LovingV). Under the type-shifting hypothesis, EnjoyingV and LovingV are taken to differ in that only the former exclusively selects for eventive complements whereas the latter does not. From our perspective, both EnjoyingV and LovingV are psychological verbs. We have no independent reason to say that the two differ in terms of lexical semantics. The reason we separate them here is because the verbs in the EnjoyingV condition have been tested in the literature as “coercion verbs” while those included in the LovingV set have not. Part of the contribution of Experiment 1 (Sect. 3) is to determine if there could be processing differences between these two sets at all that could warrant two different linguistic treatments.

Aside from the 50 triplets, we introduced 150 filler sentences (50 of them were nonsensical sentences); the whole set of the stimuli amounted to 300 sentences. The 50 triplets were split into two lists, each list containing 25 triplets of the three conditions along with half of the fillers. Each list was assigned to 20 participants, and each participant received a unique pseudo-randomization of her corresponding list.

2.1.3 Procedures

The participants were asked to rate the acceptability of each sentence on a 1–5 scale (1 = does not make sense; 5 = makes sense) and answer a multiple-choice, multiple-answer question probing possible interpretations.

2.2 Results

Results from the means (Table 2) show that the three conditions were within the acceptable range.

Repeated measures Analysis of Variance (ANOVA) revealed an effect of condition ($F(2, 72) = 32.59, p < 0.001$). Planned pairwise comparisons indicated that LovingV was rated significantly higher than both AspectualV and EnjoyingV respectively (both

Table 2 Results of the sensality rating (N = 37)

Condition	Mean	Standard deviation
AspectualV	4.13	0.75
EnjoyingV	4.31	0.52
LovingV	4.80	0.22

$ps < 0.001$).¹¹ Crucially, no difference was found between the AspectualV and EnjoyingV conditions ($p > 0.05$). In addition, we performed a reliability test on the items within each condition to evaluate their internal consistency. The reliability results showed that the items used in each condition were highly reliable (Cronbach's alpha: AspectualV = 0.92; EnjoyingV = 0.85; LovingV = 0.72). This means that, while each condition contained 50 items (sentences), the items within each condition were closely related as a group, yielding similar responses.

3 Experiment 1: Self-paced Reading

We conducted a self-paced reading experiment with a moving window paradigm to investigate the time-course of the cost underlying the processing of aspectual verbs and psychological verbs.

3.1 Method

3.1.1 Participants

Twenty-eight native speakers of American English were recruited, all between the ages of 18–30 and with normal vision and auditory acuity. None of them had history of reading disabilities.

3.1.2 Materials

The materials were adapted from the pretest questionnaire. The script contained 50 triplets, each consisting of the three conditions, and 150 filler sentences (among

¹¹We attribute the difference in rating scores between LovingVs in one hand, and the other two conditions on the other hand as a difference in familiarity induced by frequency. Verbs like *love*, *like*, *dislike* have higher frequencies of use in the language, which means that the participants are more likely to have come across these verbs in a greater variety of contexts and thus building a greater familiarity with the usage. Indeed, a closer look at the frequency distribution via Corpus of Contemporary American English suggests that LovingVs such as *loved* (42630), *liked* (32715), *disliked* (1522), *hated* (10900) have higher frequencies (in parentheses) than EnjoyingVs such as *enjoyed* (16000), *preferred* (12438), *favored* (7637), and *tolerated* (2187).

Table 3 Example of a set of experimental sentences and segmentations

Condition		Verb	Complement head	Head+1	Head+2	
AspectualV	<i>Lady Gaga</i>	<i>started</i>	<i>this CD</i>	<i>of</i>	<i>American</i>	<i>pop hits.</i>
EnjoyingV	<i>Lady Gaga</i>	<i>preferred</i>	<i>this CD</i>	<i>of</i>	<i>American</i>	<i>pop hits.</i>
LovingV	<i>Lady Gaga</i>	<i>loved</i>	<i>this CD</i>	<i>of</i>	<i>American</i>	<i>pop hits.</i>

which 100 were nonsensical). Each participant saw all the 300 sentences during the experiment. Each sentence was segmented into several windows as shown in Table 3. Our windows of interest were the Verb, Complement head, Head+1, and Head+2 regions.

The verbs used (and their frequencies) in the stimuli are the following: AspectualVs included *start* (12), *begin* (13), *finish* (12), *continue* (9), *complete* (3), *end* (1). EnjoyingVs included *enjoy* (14), *prefer* (10), *favor* (9), *tolerate* (8), *endure* (8), *resist* (1). These verbs were adopted from previous studies and have been used in Katsika et al. (2012). LovingVs included *love* (13), *like* (12), *dislike* (8), *hate* (5), *detest* (5), *approve of* (3), *be fond of* (2), *disapprove of* (1), *respect* (1). All verbs in the three conditions were matched by reaction times from an independent lexical decision study carried out in our lab (DiNardo, unpublished thesis). It showed no difference in accessing times among AspectualVs (465.77 ms), EnjoyingVs (454.46 ms), and LovingVs¹² (488 ms), all *ps* > 0.05.

3.1.3 Procedure

The stimuli were presented in black Courier New font in the center of a computer screen with a white background. The participants read the sentences segment by segment at their own pace, which allowed them to fully understand the sentences' meanings. Every trial began with a series of dash lines, with a "+" sign at the left edge of the screen, signaling the starting point of the sentence. The participants began by pressing the space bar, causing the first segment to show up. With the subsequent press, the next word appeared, and the previous segment was replaced by a set of dashes. At the end of the sentence, they were presented a statement probing either the content or the acceptability of the sentence just read to ensure full comprehension. The participants responded by pressing the "Agree" or "Disagree" key on the keyboard. A practice session was given beforehand; the participants had to reach 80% accuracy in the comprehension task before proceeding to the real trials.

¹²The only item (of the LovingV condition) not included in this lexical decision study is "*be fond of*," whose length affects RTs.

3.1.4 Data Analysis

All 28 participants recruited were taken into account in the data analysis; none was excluded. We performed a mixed model analysis, incorporating a fixed effect of *condition* (3 levels: AspectualV, EnjoyingV, LovingV) and random intercepts for *subject* and *item*. Analyses were carried out in the R statistical environment, using the *lmer* function in the *lme4* packages (Baayen et al. 2008; R Development Core Team 2014). The reading time measure was evaluated by contrasting a model including condition as the predictor against a base model without it. This contrast shows whether there is a significant effect of condition. For the pairwise comparisons, the *p*-values were corrected by Tukey tests, and the *b* values represent the unstandardized coefficients. All significant contrasts are reported.

3.2 Results

The accuracy of the comprehension task was 95.03%. Results of the reading task are reported in Table 4 and Fig. 1. A marginally significant effect of condition was found at the verb ($\chi^2(2) = 5.475, p = 0.0647$: AspectualV/EnjoyingV > LovingV) which went away at the complement head ($\chi^2(2) = 1.445, p = 0.486$). Sustained significant differences appeared instead at the Head+1 and Head+2 positions.

At the Head+1 position, there was a significant effect of condition ($\chi^2(2) = 14.315, p < 0.001$). The pairwise comparisons indicate that AspectualV engendered significantly longer reading times (RTs) than EnjoyingV ($b = 18.301, p = 0.036$) and LovingV ($b = 27.581, p < 0.001$) respectively.

The Head+2 position revealed the same pattern. A significant effect of condition was found ($\chi^2(2) = 11.197, p = 0.004$). The pairwise comparisons suggest that AspectualV engendered longer RTs than both EnjoyingV ($b = 27.172, p = 0.011$) and LovingV ($b = 27.504, p = 0.010$).

Overall, the results indicate that aspectual verbs induced longer RTs than both the *enjoy*-type and the *love*-type of psychological verbs at the two windows following the complement head, while the two types of psychological verbs did not differ from each other. These results replicate Katsika et al.'s (2012) and Utt et al. (2013) findings, and further show that psychological verbs behave as a class in terms of processing profile. We interpret these findings as suggesting that the unique cost observed for aspectual verbs, but not psychological verbs (EnjoyingV

Table 4 Results of reading times in millisecond (standard errors in parenthesis)

	Verb	Complement head	Head+1	Head+2
AspectualV	553.45 (7.30)	608.01 (8.70)	502.08 (7.13)	538.99 (9.07)
EnjoyingV	556.42 (7.26)	611.64 (9.61)	483.78 (5.55)	511.82 (7.38)
LovingV	537.93 (6.58)	598.98 (9.34)	474.50 (5.28)	511.49 (7.59)

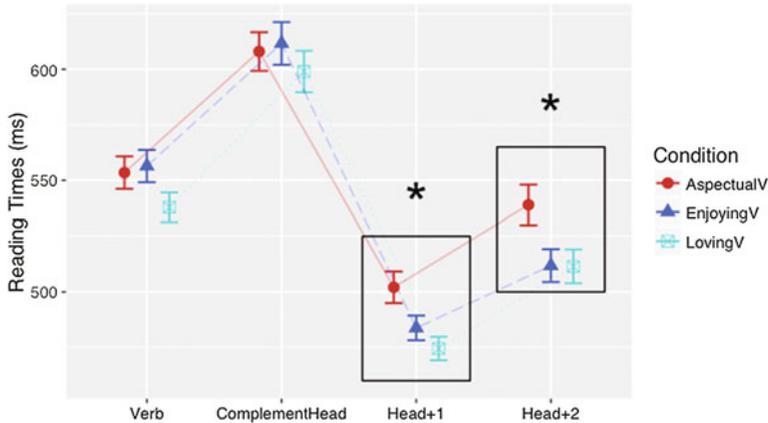


Fig. 1 Results of reading times (ms); error bars: ± 1 standard error of the means

and LovingV), is due to the specific interpretive requirements of the aspectual verb class: the determination among multiple possibilities of the dimension (e.g. *eventive*, *informational*, *spatial*) along which the structured individual associated with the complement denotation must be construed. Crucially, this determination is required; failure to resolve the ambiguity leads to failure to interpret the sentence.¹³

4 Experiment 2: fMRI

Previous neurological studies of the complement coercion effect report activity in three distinct brain regions: Wernicke’s area in Piñango et al. (2001) lesion study, ventral medial prefrontal cortex (vmPFC) in Pykkänen and McElree’s (2007) MEG study, and BA45 in Husband et al.’s (2011) fMRI experiment. Despite the discrepant results, all these studies attribute the effect to type-shifting the complement to obtain an event interpretation. Notice that these experiments are subject to the heterogeneous stimuli problem, mixing aspectual verbs, psychological verbs, and

¹³Two anonymous reviewers suggest that these results in themselves are still consistent with the type-shifting account provided that coercion is restricted to aspectual verbs. In footnote 5 above we pointed out one key problem with this solution. Here we also note that the failure of the EnjoyingV condition (composed strictly by those psychological verbs previously claimed to be coercion verbs) to behave differently from the LovingV condition represents a further argument against the type-shifting account, since the verbs in the EnjoyingV condition but not in the LovingV condition passed what had been proposed as the main diagnostic for coercion: participation in the [V + V-ing] paraphrase. What the unified behavior of the psychological verbs in the SPR shows is that this diagnostic is not restrictable to aspectual verbs (since EnjoyingVs pass it) and therefore not adequate for the characterization of the coercion effect understood in terms of mismatch and repair via type-shifting.

control verbs (e.g. *master*, *try*). Yet as mentioned, recent studies (Katsika et al. 2012; Utt et al. 2013) indicate that only aspectual verbs engender additional cost. On the other hand, the Structured Individual hypothesis accounts for the complement coercion effect as the processing of aspectual verbs. In our fMRI experiment, we aim to investigate its neural basis, expecting to find brain activity corresponding to the two hypothesized processes associated with it.

4.1 Method

4.1.1 Participants

Sixteen native speakers of American English participated this study, all between the ages of 18–30, right-handed, without reading disabilities or history of neurological disorders. The data from one participant were excluded from the analysis due to severe head movement.

4.1.2 Materials

The stimuli were the same as the self-paced reading experiment, with a different set of fillers. The 50 triplets and 150 fillers yielded 300 sentences in total. Each participant saw all 300 sentences.

4.1.3 Experimental Design

We used an event-related paradigm. The stimuli were visually presented segment-by-segment as in the self-paced reading experiment, each lasting for 500 ms. For 75% of the sentences, the participants were queried with a yes/no comprehension question, which lasted for 4000 ms. There was a 500 ms interval between the sentence-final word and the comprehension question following the sentence.

The total 300 sentences were divided into 10 runs. The stimuli were pseudo-randomized such that no successive sentences were of the same condition. Each run contained 30 sentences and lasted 5 min 33 s with the inclusion of the machine connection delay.¹⁴

¹⁴This included the time to connect the stimulus presentation software (E-Prime) and the MRI scanner.

4.1.4 Imaging Acquisition

Anatomical Measurements: The fMRI experiment was performed on a Siemens Sonata; 3T whole body MRI scanner (Erlangen, Germany). Each session began with a 3-plane localizer followed by a sagittal localizer, and an inversion recovery T1 weighted scan (TE/TR = 2.61/285 ms, matrix 192×192 , FOV = 220 mm, flip angle = 70° , bandwidth = 501 Hz/pix, 51 slices with 2.5 mm thickness). This acquisition was used to define the AC-PC (anterior and posterior commissure) line for prescription of the anatomic T1 images and functional images in the following series.

Functional measurements: During the task, we conducted event-related functional MRI using gradient echo echo-planar imaging (EPI) blood oxygenation level dependent (BOLD) contrast, with TE = 30 ms, TR = 956 ms, matrix 84×84 , FOV = 210 mm, flip angle = 62° , bandwidth = 2289 Hz/pixel, slice thickness = 2.5 mm, with 321 measurements (images per slice). The scanner was set to trigger the stimulus presentation program, which enabled the image acquisition to be synchronized with the stimulus presentation.

At the end of the functional imaging, a high-resolution 3D Magnetization Prepared Rapid Gradient Echo (MPRAGE) was used to acquire sagittal images for multi-subject registration, with TE = 2.77 ms, TR = 2530 ms, acquisition matrix 256×256 , FOV = 256 mm, bandwidth = 179 Hz/pix, flip angle = 7° , 176 slices with slice thickness = 1 mm (the fMRI data within subjects was registered to this brain volume, which was then registered across subjects into a common 3D brain space using the Yale BioImage Suite software package (Papademetris et al. 2006).

4.1.5 fMRI Data Analysis

All data were converted from Digital Imaging and Communication in Medicine (DICOM) format to analyze format using XMedCon (Nolfe et al. 2003). During the conversion process, the first 6 images at the beginning of each of the 10 functional runs were discarded to enable the signal to achieve steady-state equilibrium between radio frequency pulsing and relaxation, leaving 315 images per slice per run for analysis. Functional images were motion-corrected with the Statistical Parametric Mapping (SPM) 5 algorithm (<http://www.fil.ion.ucl.ac.uk/spm/software/spm5>) for three translational directions (x, y or z) and three possible rotations (pitch, yaw or roll). Trials with linear motion that had a displacement in excess of 1.5 mm

Table 5 Event segmentation in the fMRI experiment

Condition	Event 1 (subject + verb)	Event 2 (complement ~ sentence-final)
AspectualV	<i>Lady Gaga <u>started</u></i>	<i>this CD of American pop hits.</i>
EnjoyingV	<i>Lady Gaga <u>preferred</u></i>	<i>this CD of American pop hits.</i>
LovingV	<i>Lady Gaga <u>loved</u></i>	<i>this CD of American pop hits.</i>

or rotation in excess of 2° were rejected. All further analyses were performed using BioImage Suite (<http://bioimagesuite.org>, Papademetris et al. 2006).

Individual subject data was analyzed using a General Linear Model (GLM) on each voxel in the entire brain volume with regressors specific for each task. For each of the 3 sentence types (AspectualV, EnjoyingV, LovingV) there were two regressors for two events, as shown in Table 5. These events correspond to the two hypothesized processes induced by aspectual verbs under the SI hypothesis. **Event 1** included the onset of the subject phrase and the offset of the verb; **Event 2** included the onset of the complement until the offset of the sentence-final word. We hypothesize that the exhaustive activation of an aspectual verb's functions takes place at Event 1 when readers encounter the verb and that readers attempt to determine the dimension along which the complement is construed as a structured individual at Event 2 while facing the dimension ambiguity in sentences with aspectual verbs.

The resulting beta images for each task were spatially smoothed with a 6 mm Gaussian kernel to account for variations in the location of activation across subjects. The output maps were normalized beta-maps, which were in the acquired space ($2.5 \text{ mm} \times 2.5 \text{ mm} \times 2.5 \text{ mm}$).

To take these data into a common reference space, three registrations were calculated within the Yale BioImage Suite software package. The first registration performed a linear registration between the individual subject raw functional image and that subject's 2D anatomical image. The 2D anatomical image was then linearly registered to the individual's 3D anatomical image. The 3D differs from the 2D in that it has a $1 \times 1 \times 1 \text{ mm}$ resolution whereas the 2D z-dimension is set by slice-thickness and its x-y dimensions are set by voxel size. Finally, a non-linear registration was computed between the individual 3D anatomical image and a reference 3D image. The reference brain used was the Colin27 Brain (Holmes et al. 1998) in Montreal Neurological Institute (MNI) space (Evans et al. 1992). All three registrations were applied sequentially to the individual normalized beta-maps to bring all data into the common reference space.

Data were corrected for multiple comparisons by spatial extent of contiguous suprathresholded individual voxels at an experiment-wise $p < 0.05$. In a Monte Carlo simulation within the AFNI software package and using a smoothing kernel of 6 mm and a connection radius of 4.33 mm on $2.5 \text{ mm} \times 2.5 \text{ mm} \times 2.5 \text{ mm}$ voxels, it was determined that an activation volume of 183 original voxels (4953 microliters) satisfied the $p < 0.05$ threshold. Clusters were created for each of the subtractions. Each cluster was identified with a region label, and then associated with additional numeral labels corresponding to Brodmann areas.

Table 6 Mean response times (ms) of the comprehension questions

Condition	Mean	Standard error (se)
AspectualV	1684.18	25.06
EnjoyingV	1722.11	25.08
LovingV	1651.49	24.09

4.2 Results

4.2.1 Behavioral Results

The overall accuracy of the comprehension task was 88.6%. The response times (RTs) for the questions are shown in Table 6. Results of repeated measures ANOVA (trials with no response were excluded) revealed no significant effect of conditions in RTs ($F(2, 28) = 2.786, p = 0.08$).

Fig. 2 AspectualV > EnjoyingV at Event 1 (Subject + Verb): activations in left BA 40 (Wernicke’s area), bilateral BA7, bilateral BA 6/24, and bilateral primary sensory areas

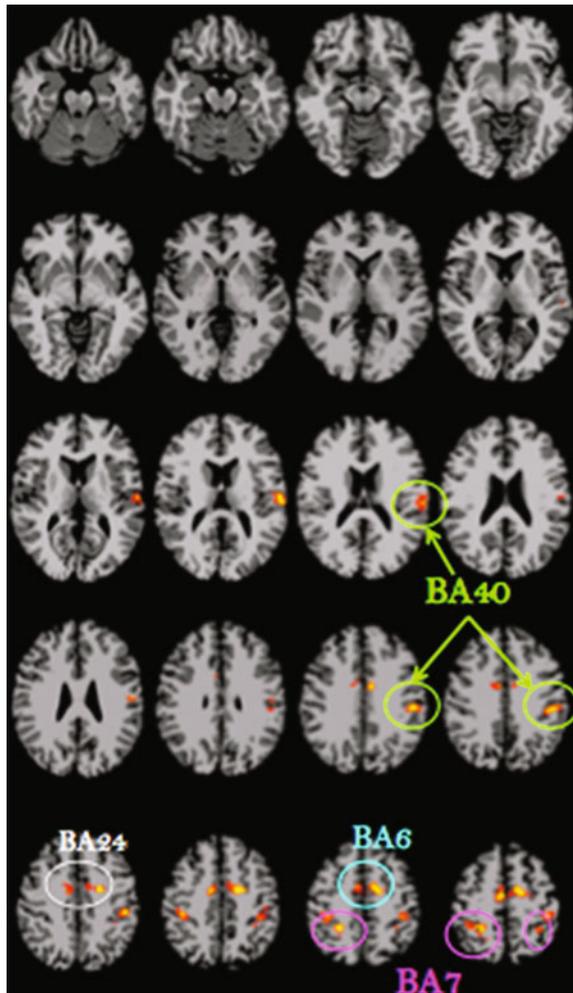


Fig. 3 AspectualV > LovingV at Event 2 (Complement ~ S-final): activations in LIFC (BA44, 45, 47), left insula, right BA8, bilateral BA6, right IF cortex, and left primary visual cortex

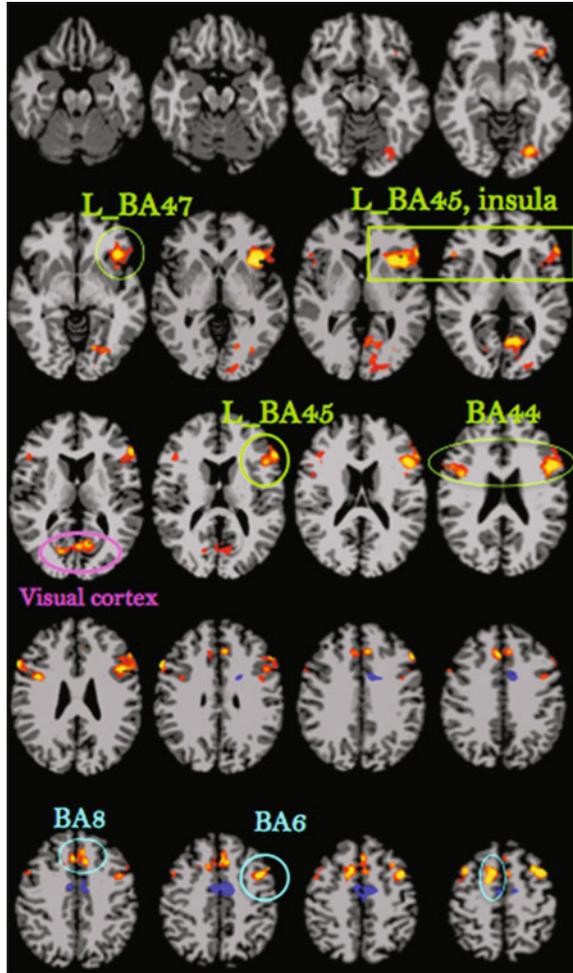


Table 7 Summary of the results of the fMRI experiment (significant activations)

	Event 1 (subject + Verb)	Event 2 (complement ~ S-final)
AspectualV — EnjoyingV (Fig. 2)	AspectualV recruited: <ul style="list-style-type: none"> • Left BA 40 (Wernicke’s area) • Bilateral BA 7, 6, 24 • Bilateral primary sensory cortex 	EnjoyingV recruited more left BA6, 24 and their right counterparts in upper brain.
AspectualV —LovingV (Fig. 3)	LovingV recruited more left posterior occipital regions (BA7, 18, 19, part of BA39), and right BA 44/45.	AspectualV recruited: <ul style="list-style-type: none"> • LIF cortex (BA 44, 45, 47), left insula • Bilateral BA6, right BA8 • Left primary visual cortex • Right inferior frontal cortex
EnjoyingV —LovingV	LovingV recruited more right BA7 and right BA39	EnjoyingV recruited more bilateral BA44, 45, 47, insula, and bilateral BA6, 8

Table 8 The differentially active regions of the AspectualV—EnjoyingV at Event 1_(Subject+Verb). L = left, R = right, AntCingulate = anterior cingulate, PrimSensory = primary sensory cortex, Prim_Motor = primary motor cortex

Region	Volume (mm ³)	Mean T-Value	Max T-Value	MNICoord Max
L BA40	2168	2.52033	3.55617	-48, -27, 39
L BA7	2761	2.6424	4.39206	-30, -54, 69
Medial_BA6_AntCingulate	7735	2.53566	4.9745	-18, -6, 45
R_BA7_PrimSensory	7377	2.61813	5.71615	27, -45, 51
L PrimSensorv PrimMotor	4081	2.44217	4.02101	-60, -15, 12

Table 9 The differentially active regions of the AspectualV—LovingV at Event 2_(Complement ~S-final) L = left, R = right, IFG = inferior frontal gyrus, AntCingulate = anterior cingulate

Region	Volume (mm ³)	Mean T-Value	Max T-Value	MNICoord Max
L Insula	1473	3.28132	6.8164	-36, 18, 0
L IFG BA47 BA45 BA44	11565	2.75671	6.73777	-37, 18, 0
L lateral BA6	3718	2.83759	5.80195	-39, 9, 51
R_AntCingulate	248	2.63028	3.68964	-6, 30, 30
R IFG BA44 BA45	3135	2.57169	4.19239	60, 18, 27
R medial BA6 BA8	5647	2.70576	5.40665	15, 0, 54
R lateral BA6 BA8	1080	2.53198	5.34182	39, 12, 54
Medial BA6 AntCingulate	4958	-2.59486	-4.74066	-18, -15, 78
PrimarvVisual BA18 BA19	6148	2.54813	4.2163	-12, -63, 6

4.2.2 Imaging Results

The imaging results showed that, at **Event 1** (Subject + Verb), AspectualV preferentially recruited Wernicke's area (BA40), bilateral BA7, 6, 24, and primary sensory areas over EnjoyingV (Fig. 2). At **Event 2** (Complement ~ Sentence-final), AspectualV preferentially recruited the left inferior frontal cortex (LIFC), including BA44, 45, 47, and left insula, as well as bilateral BA6, right BA8, right inferior frontal cortex, and primary visual cortex over the control LovingV (Fig. 3). These results are summarized in Tables 7, 8 and 9.

Overall, the fMRI results reveal that at Event 1 (Subject + Verb) AspectualV shows preferential recruitment in Wernicke's area (BA 40) over EnjoyingV (Fig. 2) whereas at Event 2 (complement phrase), preferential recruitment for AspectualV (over LovingV) shifts to left inferior frontal cortex (Fig. 3).

5 Discussion

Experiment 1 (self-paced reading) showed that the AspectualV condition engendered longer reading times than psychological verbs (EnjoyingV and LovingV) during real-time comprehension after the complement head was encountered, thus replicating and expanding on Katsika et al.'s (2012) and Utt et al.'s (2013) findings. These observations call into question our traditional understanding of the “complement coercion” phenomenon. The fact that aspectual and psychological verbs show not only distinct linguistic behaviors but also computationally distinct processing profiles suggests that the set of “coercion verbs” studied in previous work collapses at least two semantically distinct classes of verbs. These observations are in principle inconsistent with the type-shifting hypothesis, which predicts processing cost at least for those psychological verbs that fall in the coercion set—a prediction that is not supported by the results.

In order to account for this pattern of results, we propose the Structured Individual hypothesis (based on the Structured Individual analysis for aspectual verbs) whereby the processing cost associated with aspectual verbs results from (A) exhaustively retrieving the possible functions stored in the verb, and (B) the resolution of dimension ambiguity (e.g. *spatial*, *temporal*, *informational*, *eventive*) that is required to interpret the complement and consequently, the full sentence. According to this hypothesis, the dimension is identified when the complement head denoting a structured individual is encountered and the dimension along which this structure must be mapped onto the axis is determined.¹⁵

We already find initial support for this general approach in previously puzzling findings. Traxler et al. (2005) show that a prior context sentence that either contained the same coercion verb or explicitly mentioned the event structure attenuated the cost in the following target sentence.

- (15) Context: *The student started/read a book...*
 Target: *Before he started/read the book...*

In their eye-tracking experiment, when either “*started a book*” or “*read a book*” is given in the context sentence, there was no difference between “*started the book*” and “*read the book*” in the target sentences. The Structured Individual

¹⁵Crucially, the dimension ambiguity resolution engendered by aspectual verbs is distinct from inferencing the specific activity associated with the complement in the agentive interpretation in the eventive dimension. Indeed, Frisson and McElree (2008) argue that the complement coercion effect is not modulated by competition or selection/retrieval of a particular event. In their eye-movement experiment, strongly preferred coercion sentences (*The teenager began the novel...*) did not differ from weakly preferred ones (*The waitress started the coffee...*). This suggests that the coercion cost is not modulated by the preference of interpretation or the number of different interpretations. In fact, both the strongly and weakly preferred sentences in their experiment involve aspectual verbs and animate subjects. Under the SI hypothesis, both sentences give rise to an ambiguity along different dimensions (e.g. informational, temporal, or eventive), and therefore cost emerges independent of the preference, as it is observed.

(SI) hypothesis captures this naturally: because the proper interpretation among multiple ones is contextually determined, a context sentence that biases the interpretation towards a certain dimension (likely the eventive dimension in this case) resolves the dimension ambiguity associated with aspectual verbs and therefore attenuates the cost.

A seemingly problematic finding for the SI hypothesis is Traxler et al.'s (2002) results. They show that, following coercion verbs (including both aspectual and psychological), entity-denoting complements (*started the puzzle*) induced more processing cost than event-denoting complements (*started the fight*). This may seem to go against the SI hypothesis, which states that any sentence containing aspectual verbs (including those with eventive complements) will require both exhaustive dimension-function retrieval and dimension-ambiguity resolution (the two potential sources of cost). The SI hypothesis accounts for this reported difference as an attenuation of the effect brought about by a pre-determination of the intended dimension along which the structured individual must be construed. Recall that the objective of interpreting AspectualV sentences is to decide the exact dimension along which the structured individual is construed. In cases like *started the fight*, the event-denoting complement (“*fight*”) determines, or strongly biases towards, the eventive dimension along which a structured individual is construed. Because the dimension along which the structured individual must be construed is already determined by the eventive complement, the effort of resolving the dimension ambiguity (process B) is decreased, thus the observable cost of the composition is attenuated.

We focus now on Experiment 2 (fMRI), which shows that aspectual verbs induced preferential recruitment of the left posterior superior temporo-parietal cortex (i.e. Wernicke's area, BA40) when readers encountered the verb and the left inferior frontal cortex (i.e. Broca's area) when readers encountered the complement, over each of the psychological conditions (*enjoy*-type and *love*-type respectively). The Structured Individual hypothesis captures these patterns as follows: At **Event 1, BA40** (green circles in Fig. 2) reflects the exhaustive activation of the aspectual verb's dimension functions. The reason that aspectual verbs would engage this cortical region above and beyond the engagement induced by the EnjoyingV condition is presumably that only the former must encode discrete, pre-specified, functions. This encoding represents a measure of complexity and therefore of cost. This interpretation of Wernicke's area is consistent with previous work showing not only that it is involved in the retrieval of lexico-semantic representations (Badre et al. 2005; Binder et al. 2009; Damasio et al. 1996; Hickok and Poeppel 2004, 2007; Humphries et al. 2007; Lau et al. 2008) but also as Shapiro et al. (1993) show through a Broca's versus Wernicke's comparison, that this cortical region is specifically involved in exhaustive retrieval of lexical items during comprehension.

The activation of the precuneus (BA 7) (pink circles in Fig. 2), which was also preferentially recruited for AspectualV at Event 1, although not initially expected given previous neurolinguistic work on complement coercion, is completely consistent with our view of the fundamental meaning of aspectual verbs. This cortical region is reported to support spatio-temporal tasks such as processing visual-spatial

information in perception and memory, especially spatial representation of sequential movements (Cavanna and Trimble 2006; Fletcher et al. 1995; Wallentin et al. 2008). We therefore interpret the recruitment of BA7 for aspectual verbs over psychological verbs as reflecting axial conceptualization associated with the set of dimension functions, one of which will ultimately map the structured individual onto an axis. The activation of BA40 and BA7 combined is correlated with dimension function retrieval, which in the absence of a complement is still underdetermined. These observations are clearly grounded in the basic principles of the SI hypothesis, and in this way support it. By contrast, they are not naturally connectable to the type-shifting approach, which by its very nature is expected to target only one cortical region, and cannot be naturally connected to spatial conceptualization.

The neurological patterns discussed in the paper are those that can be clearly interpreted in the context of the comparison between the Structured Individual hypothesis and the Type-Shifting hypothesis. To the extent that EnjoyingV shows different patterns from AspectualV, we interpret the results as suggesting that the two sets of verbs do not belong to a unified class.

At **Event 2**, the LIF cortex and insula activations (green in Fig. 3) are taken to reflect the process of determining the dimension along which the complement denotation is structured. This is not an unexpected activation pattern given that the LIF cortex has in the past been reported to support certain kinds of ambiguity resolution (e.g. Badre et al. 2005; Krain et al. 2006; Lau et al. 2008; Rodd et al. 2010). In particular, several studies have reported that the insula is involved in processing of time. For instance, in an fMRI study using prismatic adaptation (PA)¹⁶, Magnani et al. (2014) had participants perform a time reproduction task, in which they indicated the time intervals as perceived. The PA-induced rightward aftereffect resulted in an overestimation of time intervals whereas the leftward aftereffect resulted in an underestimation of time intervals. Their imaging results reveal that the left anterior insula and left superior frontal gyrus showed increased activity *after* versus *before* PA. They hence suggest that these regions are involved in spatial manipulation of the representation of time. Connecting those results with the ones reported here, we propose that the insula recruitment for aspectual verbs is likely to reflect a more general kind of structural configuration, more specifically the precedence relation on the axis along some dimension, including the temporal dimension.

For Event 2 (as for Event 1) the premotor area or the supplementary motor area (BA6) was recruited preferentially for AspectualV (light-blue circle in Fig. 3). This area is reported to be involved in action planning or event sequencing (e.g. Crozier 1999), action simulation, the generation of ordinally structured sequences (Stadler et al. 2011), and the updating of spatial information (Tanaka et al. 2005).

¹⁶During PA, participants are asked to wear prismatic lenses, which shift their visual field toward one side of space, and to point at visual targets. Due to the prismatic lenses, the visual targets are shifted relative to their actual locations as perceived. The participants have to adjust to the visual twist so as to point to the visual targets precisely.

We propose that the BA6 activation associated with the AspectualV condition reflects the sequential action planning along the eventive dimension.

Finally, we take the activity in the visual cortex (purple in Fig. 3) to reflect components not specifically relevant to the linguistic mechanisms in question, but mechanisms connected to attention, generalized task difficulty (Chen et al. 2008; Tan et al. 2001), or written word recognition (see review in Price, 2012). We leave it for future research to further clarify the connection between those presumed general processes and the specific meaning composition mechanisms discussed here.

Our fMRI results replicate Piñango et al. (2001) and Husband et al.'s (2011) findings. Whereas Piñango et al. (2001) report that Wernicke's aphasics (with damage to the left posterior superior temporo-parietal cortex) have difficulty comprehending sentences involving complement coercion, Husband et al. (2011) show that BA45 in LIF cortex (Broca's area) is preferentially involved in the implementation of sentences demanding complement coercion (that is, to the extent that a large portion of their stimuli sentences used aspectual verbs). Given their corresponding experimental designs, it was not possible in those studies to elucidate *when* during the course of comprehension, these respective cortical regions were maximally recruited. Experiment 2 presented here resolves this question by showing that Wernicke's area is preferentially recruited during the unfolding of the subject + aspectual verb composition whereas LIF cortex is preferentially recruited later in the comprehension process, once the complement head (indicating the presence of a structured individual) is retrieved. And again, this multiplicity of regions finds coherence in the Structured Individual hypothesis while it is left unexplained in the type-shifting approach, which by definition requires the activation to be localized to one cortical region.

Regarding the psychological verb conditions, as we have seen, the fMRI results show a "split" behavior such that for Event 1, only *EnjoyingV* shows a difference in preferential recruitment with respect to AspectualV; and for Event 2, only *LovingV* shows a difference in preferential recruitment with respect to AspectualV. We believe that this difference in behavior is rooted in independent but potentially interacting factors which the present results can only begin to tease apart. So, what we offer here is a conjecture whose examination we leave for future work.

As mentioned in the Introduction section, we do not predict a linguistic difference between the *EnjoyingV* and *LovingV* conditions. Consistent with previous proposals, verbs in both conditions are all equally expected to select for a target of emotion (Pesetsky 1995: 55, 96; Levin 1993; Katsika et al. 2012), thus predicting unified psycholinguistic behavior, which the self-paced reading results reported here support.

Moreover, we note that it is only with respect to *composition with the complement* (Event 2) that our analysis predicts preferential recruitment for aspectual

verbs above and beyond that for the two psychological verb conditions. With respect to *composition with the subject* (Event 1), our analysis *in principle* makes no such prediction. The reason for this is that like aspectual verbs, psychological verbs also select for both entity-denoting and event-denoting complements. It is possible that this complexity in lexical encoding gives rise to comparable degree of neurological recruitment in Wernicke's area during lexical retrieval of the verbs.

To the extent that we found a difference at Event 1 between the AspectualV and LovingV sets on one hand and the EnjoyingV set on the other hand (AspectualV = LovingV > EnjoyingV), we look for the cause in the interpretive biases of this latter subset of psychological verbs. Indeed, what unites the verbs in the EnjoyingV set is that even though they may select for entity-denoting complements, the complement is preferentially interpreted eventively i.e. the preferential interpretation of *enjoy the book* is as *enjoy reading/writing the book*.¹⁷ This contrasts with the LovingV set in that these verbs allow for such a paraphrase, but not necessarily preferentially. So, *love the book*, can be understood as *love reading the book*, but does not have to be; it can also be understood as *love the story in the book* (state). If this eventive bias in interpretation for EnjoyingV were to impact lexical retrieval (by allowing the event-denoting possibility to be considered first by the processor), it could make the EnjoyingV set less ambiguous and therefore less taxing on Wernicke's area than the LovingV set. This could, in turn, result in a greater difference between AspectualV and EnjoyingV over AspectualV and LovingV at Event 1.

Finally, regarding Event 2, our hypothesis predicts AspectualV > {EnjoyingV = LovingV}, a prediction that holds only for LovingV. In line with our conjecture, we observe that EnjoyingV condition, constructed with an event-selecting bias, would be tapping a similar conceptual representation as the AspectualV set, which disambiguates along the eventive dimension.¹⁸ Both conditions, AspectualV and EnjoyingV, would then recruit possibly overlapping cortical regions. Importantly, this would take place during Event 2, the segment when the decision in favor of the eventive dimension would have been made—a reading that, crucially, is not equally salient in the LovingV condition.

All this said, what is most important for our analysis is that the traditional coercion verb set shows a split in neurological recruitment of Wernicke's area (AspectualV > EnjoyingV) at Event 1. And this split is consistent with the linguistic distinction proposed here.

¹⁷A closer look at their distribution of the specific experimental V + Complement in COCA (Corpus of Contemporary American English) shows that the LovingV set appears with greater frequency for the entity-denoting complements used in the stimuli than the EnjoyingV set (LovingV: 41; EnjoyingV: 7).

¹⁸Recall that the eventive dimension in the agentive interpretation is a salient dimension that was tapped in the AspectualV condition.

6 Conclusion

This study represents an implementation of the Structured Individual analysis, working out its psychological and neurological viability in the form of the Structured Individual hypothesis; a hypothesis that captures a psycho- and neurolinguistic distinction between aspectual verbs and psychological verbs in a linguistically principled manner, and invokes independently motivated processing mechanisms to capture their behavior. Our findings suggest that the complement coercion effect is better understood as involving the real-time composition of aspectual verbs rather than involving a special semantic operation such as type-shifting. In this way, this kind of hypothesis represents a manifestation of an approach to meaning composition which connects functional application to concept composition through lexicalization, all along grounded in fundamental principles of conceptual structure.

Acknowledgements Research was funded by NSF-BCS grant BCS-0643266 to Maria Mercedes Piñango, and NSF-INSPIRE Grant CCF-1248100 to Maria Mercedes Piñango, Ashwini Deo, Todd Constable, and Mokshay Madiman. We thank Edgar Zurif, Ray Jackendoff, James Pustejovsky for very useful feedback, and audiences at Tsing Hua National University, Chung Cheng University (Taiwan 2013), the CUNY conference on Sentence Processing 2014, and the Annual Cognitive Science Meeting 2014. All errors remain our own.

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? a critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796.
- Badre, D., Poldrack, R. A., Juliana Paré-Blagoev, E., Inslar, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, 47(6), 907–918.
- Baggio, G., Choma, T., van Lambalgen, M., & Hagoort, P. (2010). Coercion and compositionality. *Journal of Cognitive Neuroscience*, 22(9), 2131–2140.
- Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, 129(3), 564–583.
- Chen, Y., Martinez-Conde, S., Macknik, S. L., Bereshpolova, Y., Swadlow, H. A., & Alonso, J.-M. (2008). Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. *Nature Neuroscience*, 11(8), 974–982.
- Crozier, S., Sirigu, A., Lehericy, S., van de Moortele, P. -F., Pillon, B., Grafman, J., et al. (1999). Distinct prefrontal activations in processing sequence at the sentence and script level: An fMRI study. *Neuropsychologia*, 37(13), 1469–1476.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, 380, 499–505.
- De Swart, H. (1998). Aspect shift and coercion. *Natural Language & Linguistic Theory*, 16(2), 347–385.

- Deo, A. S., & Pinango, M. M. (2011). Quantification and context in measure adverbs. In *Proceedings of SALT* (Vol. 21).
- Deo, A., Francez, I., & Koontz-Garboden, A. (2013). From change to value difference. In *Proceeding of Semantics and Linguistic Theory.*, 23, 97–115.
- Evans, A. C., Marrett, S., Neelin, P., Collins, L., Worsley, K., Dai, W., et al. (1992). Anatomical mapping of functional activation in stereotactic coordinate space. *Neuroimage*, 1(1), 43–53.
- Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: processing multiple meanings vs. multiple senses. *Journal of memory and language* 29(2), 181–200.
- Fletcher, P. C., Frith, C. D., Baker, S. C., Shallice, T., Frackowiak, R. S. J., & Dolan, R. J. (1995). The mind's eye—precuneus activation in memory-related imagery. *Neuroimage*, 2(3), 195–200.
- Frisson, S., & McElree, B. (2008). Complement coercion is not modulated by competition: Evidence from eye movements. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 34(1), 1.
- Gawron, Jean Mark. 2009. The lexical semantics of extent verbs. Unpublished ms. San Diego State University.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1), 67–99.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402.
- Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of Computer Assisted Tomography*, 22(2), 324–333.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: an fMRI study. *Neuroimage*, 36(3), 924–932.
- Husband, E. M., Kelly, L. A., & Zhu, D. C. (2011). Using complement coercion to understand the neural basis of semantic composition: evidence from an fMRI study. *Journal of Cognitive Neuroscience*, 23(11), 3254–3266.
- Jackendoff, R. (1992). Parts and boundaries. In B. Levin & S. Pinker (Eds.), *Lexical and conceptual semantics* (pp. 9–45). Cambridge, MA: Blackwell.
- Jackendoff, R. (1997). *The architecture of the language faculty* (Vol. 28). MIT Press.
- Katsika, A., Braze, D., Deo, A., & Piñango, M. M. (2012). Complement coercion. *The Mental Lexicon*, 7(1), 58–76.
- Krain, A. L., Wilson, A. M., Robert Arbuckle, F., Castellanos, X., & Milham, M. P. (2006). Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making. *NeuroImage*, 32, 477–484.
- Krifka, M. (1998). The origins of telicity. In Rothstein, S (Ed.), *Events and grammar* (pp. 197–235). Dordrecht: Kluwer.
- Lapata, M., Keller, F., & Scheeper, C. (2003). Intra-sentential context effects on the interpretation of logical metonymy. *Cognitive Science: A Multidisciplinary Journal*, 27(4), 649–668.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago Press.
- Magnani, B., Frassinetti, F., Ditye, T., Oliveri, M., Costantini, M., & Walsh, V. (2014). Left insular cortex and left SFG underlie prismatic adaptation effects on time perception: Evidence from fMRI. *NeuroImage*, 92, 340–348.
- McElree, B., Pylikäinen, L., Pickering, M. J., & Traxler, M. J. (2006). A time course analysis of enriched composition. *Psychonomic Bulletin & Review*, 13(1), 53–59.
- McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched composition. *Cognition*, 78(1), B17–B25.
- Michaelis, L. A. (2005). *Entity and event coercion in a symbolic theory of syntax* (pp. 45–88). Construction Grammars: Cognitive Grounding and Theoretical Extensions.

- Nolfe, E., Voet, T., Jacobs, F., Dierckx, R., & Lemahieu, I. (2003). XMedCon: An open-source medical image conversion toolkit. *European Journal of Nuclear Medicine*, 30 (Suppl. 2), S246. <http://xmedcon.sourceforge.net>.
- Papademetris, X., Jackowski, M. P., Rajeevan, N., DiStasio, M., Hirohito Okuda, R., Constable, T., et al. (2006). BioImage Suite: An integrated medical image analysis suite: An update. *The Insight Journal*, 2006, 209.
- Partee, B. (1987). Noun phrase interpretation and type-shifting principles. *Studies in discourse representation theory and the theory of generalized quantifiers*, 8, 115–143.
- Partee, B., & Rooth, M. (1983). Generalized conjunction and type ambiguity. *Formal semantics: the essential readings* (pp. 334–356).
- Pesetsky, D., (1995). *Zero syntax: Experiencers and cascades*. MIT Press.
- Pickering, M. J., McElree, B., Frisson, S., Chen, L., & Traxler, M. J. (2006). Underspecification and aspectual coercion. *Discourse Processes*, 42(2), 131–155.
- Pickering, M. J., McElree, B., & Traxler, M. J. (2005). The difficulty of coercion: A response to de Almeida. *Brain and Language*, 93(1), 1–9.
- Piñango, M. M., & Deo, A. (2015). Reanalyzing the complement coercion effect through a generalized lexical semantics for aspectual verbs. *Journal of Semantics*. Advance Access published April 22, 2015, doi:10.1093/jos/ffv003.
- Piñango, M. M., & Zurif, E. B. (2001). Semantic operations in aphasic comprehension: implications for the cortical organization of language. *Brain and Language*, 79(2), 297–308.
- Piñango, M. M., Winnick, A., Ullah, R., & Zurif, E. (2006). Time-course of semantic composition: The case of aspectual coercion. *Journal of Psycholinguistic Research*, 35(3), 233–244.
- Piñango, M. M., Zurif, E. B., & Jackendoff, R. (1999). Real-time processing implications of enriched composition at the syntax–semantics interface. *Journal of Psycholinguistic Research*, 28(4), 395–414.
- Price, C. J. (2012). A review and synthesis of the first 20 years of pet and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, 62(2), 816–847.
- Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
- Pylkkänen, L., & McElree, B. (2006). The syntax-semantics interface: on-line composition of sentence meaning. In Traxler, M. J., & Gernsbacher, M. A. (Eds.), *Handbook of psycholinguistics*. New York: Elsevier.
- Pylkkänen, L., Martin, A. E., McElree, B., & Smart, A. (2009). The anterior midline field: Coercion of decision making? *Brain and Language*, 108(3), 184–190.
- Pylkkänen, L. (2008). Mismatching meanings in brain and behavior. *Language and Linguistics Compass*, 2(4), 712–738.
- Pylkkänen, L., & McElree, B. (2007). An MEG study of silent meaning. *Journal of Cognitive Neuroscience*, 19(11), 1905–1921.
- R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Rodd, J. M., Longe, O. A., Randall, B., & Tyler, L. K. (2010). The functional organisation of the fronto-temporal language system: evidence from syntactic and semantic ambiguity. *Neuropsychologia*, 48(5), 1324–1335.
- Scheepers, C., Keller, F., & Lapata, M. (2008). Evidence for serial coercion: A time course analysis using the visual-world paradigm. *Cognitive Psychology*, 56, 1–29.
- Scheepers, C., Mohr, S., Keller, F., Lapata, M. (2004) The cost of enriched composition: eye-movement evidence from German. In *Paper Presented at the CUNY Conference on Human Sentence Processing*.
- Shapiro, L. P., Gordon, B., Hack, N., & Killackey, J. (1993). Verb-argument structure processing in complex sentences in broca's and wernicke's aphasia. *Brain and Language*, 45(3), 423–447.
- Shapiro, L. P., Zurif, E. B., & Grimshaw, J. (1989). Verb processing during sentence comprehension: Contextual impenetrability. *Journal of Psycholinguistic Research*, 18(2), 223–243.

- Stadler, W., Schubotz, R. I., Yves Von Cramon, D., Springer, A., Graf, M., & Prinz, W. (2011). Predicting and memorizing observed action: Differential premotor cortex involvement. *Human Brain Mapping, 32*(5), 677–687.
- Studholme, C., Novotny, E. J., Zupal, G. I., & Duncan, J. S. (2001). Estimating tissue deformation between functional images induced by intracranial electrode implantation using anatomical MRI. *Neuroimage, 13*, 561–576.
- Tanaka, S., Honda, M., & Sadato, N. (2005). Modality-specific cognitive function of medial and lateral human Brodmann area 6. *The Journal of neuroscience, 25*(2), 496–501.
- Tan, L. H., Feng, C.-M., Fox, P. T., & Gao, J.-H. (2001). An fMRI study with written Chinese. *NeuroReport, 12*(1), 83–88.
- Todorova, M., Straub, K., Badecker, W., & Frank, R. (2000). Aspectual coercion and the online computation of sentential aspect. In *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*. 3–8.
- Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: evidence from eye-movements and self-paced reading. *Journal of Memory and Language, 47*, 530–547.
- Traxler, M. J., Williams, R. S., McElree, B., & Pickering, M. J. (2005). Context effects in coercion: Evidence from eye movements. *Journal of Memory and Language, 53*, 1–25.
- Utt, J., Lenci, A., Padó, S., & Zarcone, A. (2013). The curious case of metonymic verbs: A distributional characterization. In *Proceedings of the iwcs 2013 workshop towards a formal distributional semantics*.
- Verkuyl, H., & Zwarts, J. (1992). Time and space in conceptual and logical semantics: The notion of path. *Linguistics, 30*(3), 483–512.
- Wallentin, M., Weed, E., Østergaard, L., Mouridsen, K., & Roepstorff, A. (2008). Accessing the mental space—spatial working memory processes for language and vision overlap in precuneus. *Human Brain Mapping, 29*(5), 524–532.
- Wiese, H., & Maling, J. (2005). Beers, kaffi, and schnaps: Different grammatical options for restaurant talk coercions in three Germanic languages. *Journal of Germanic Linguistics, 17*(1), 1–38.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Conceptual Combination, Property Inclusion, and the Aristotelian-Thomistic View of Concepts

Christina L. Gagné, Thomas L. Spalding and Matthew Kostelecky

Abstract Understanding how properties are extended to combined concepts is critical to theories of concepts. In human judgments, properties true of a noun (ducks have webbed feet) become less true when that noun is modified (baby ducks have webbed feet), while properties false of a noun (candles have teeth) become less false when that noun is modified (purple candles have teeth). These modification and inverse modification effects have been shown to be extremely robust. Gagné and Spalding (2011, 2014b; Spalding and Gagné 2015) have argued that these effects are driven by expectation of contrast. The current experiment shows that, as expected, the modification and inverse modification effects are unaffected by the normative force with which a property is predicated of the head noun, supporting the expected contrast explanation. The results are discussed with respect to an Aristotelian-Thomistic approach to concepts (Spalding and Gagné 2013).

1 Overview

The human conceptual and language systems are compositional and highly productive. One important mechanism underlying this compositional and productive character is nominal compounding. Nominal compounding involves the combination of two or more free morphemes to form a noun. This process can lead to the construction of new concepts in the conceptual system and new words in the mental lexicon. Consequently, the process provides insight into both word formation and conceptual combination. Combined concepts are typically expressed as either

C.L. Gagné (✉) · T.L. Spalding
Department of Psychology, University of Alberta, Edmonton, Canada
e-mail: cgagne@ualberta.ca

T.L. Spalding
e-mail: spalding@ualberta.ca

M. Kostelecky
Department of Philosophy, University of Alberta, Edmonton, Canada
e-mail: kostelec@ualberta.ca

modifier-noun phrases (e.g., *mountain magazine*) or compound words (e.g., *snowball*) and these linguistic expressions are among the most basic examples of linguistic productivity in that they consist of the combination of two morphological constituents. In this chapter, we provide a brief overview of some psycholinguistic and cognitive research that identifies factors that influence the processing of compound words and modifier-noun phrases (which are generally either noun-noun or adjective-noun constructions). We then turn to the question of how the meanings of these linguistic expressions are elaborated, and present recent research on property verification. We discuss implications of these results for human concepts research, and end by suggesting that a view of human cognition derived from Aristotle and Thomas Aquinas could be helpful in understanding recent research on concepts.

2 Background

Concepts can often be paraphrased by expressing an underlying implicit relation between the concepts (*flu virus* is a virus that causes flu). Previous research has found that relational information plays a central role in the processing of combined concepts (for an overview, see Gagné and Spalding 2014a). The interpretation of combined concepts involves selecting the appropriate relational structure, and factors associated with the use of the relations affect the processing of the combined concept (e.g., combined concepts that share a recently used relational structure and a constituent are processed more easily than those that use a different relational structure, and these effects are true of both novel and known combinations; see Gagné and Spalding 2014a). However, such relational interpretations provide only the gist (i.e., a paraphrase) of the combined concept's meaning (e.g., that a *dog house* is a house for dogs) and do not exhaust our knowledge of the combined concept. Where does the rest of the meaning come from? For example, imagine that someone hears, for the first time, the phrase *dog house* and that they know the meaning of *dog* and *house*. Now, suppose we ask this learner, do doghouses have doorknobs? There are several interesting points to note. First, the gist interpretation *house for a dog*, does not, in and of itself, give the answer to the question. Second, the gist is not unrelated or irrelevant to answering the question. The person might reason that because dogs do not have hands, a doorknob would not be likely. Clearly, though, this is an inference based on the gist, not something that is inherently true of the combination of the concepts *dog* and *house*. Third, to be more explicit, note that *dog* and *house* could also be combined to form a gist something like *house that has a dog*, and this type of doghouse would be highly likely to have doorknobs. The key point for the current chapter is that answering the question requires an inference about the specific combination (and gist).

Our primary research question concerns how people make judgments about whether or not particular properties are true for a given concept. In terms of empirical evidence, we will focus on property verification in combined concepts, which refers to how people determine whether a particular characteristic (e.g., “is heavy”) is true

of a modified concept (e.g., *clay machine*). In addressing this question, we apply an approach in which a relation-based interpretation of the combined concept (Gagné and Shoben 1997; Spalding et al. 2010), along with specific knowledge about the constituent concepts and meta-knowledge, is used in subsequent reasoning processes to infer the content of the combined concept (Gagné and Spalding 2011, 2014b; Spalding and Gagné 2015 ; see also Connolly et al. 2007).

We should note that this “relations and reasoning” approach contrasts with many theories of conceptual combination, which tend to assume that the interpretation of a combined concept is the construction of a set of properties true of the combined concept. These properties are assumed mostly to be drawn from the constituent concepts, with some drawn from experience with the combined concept itself (see, e.g., the selective modification model of Osherson and Smith 1981; Smith and Osherson 1984; Smith et al. 1988 or schema based models such as Murphy 1988, 1990; Wisniewski 1997 or prototype combination models such as Hampton 1987, 1991). In turn, these more feature-based notions of conceptual combination tend to assume a kind of implicit “container” metaphor of concepts (see, e.g., Laurence and Margolis 1999, for a discussion of the container metaphor). Such theories tend to assume that properties are contained in the concepts (or, alternatively, that the concept just is the set of properties, or that the concept is made of the properties), and that property verification is, therefore, a matter of somehow inspecting the concept to see if the property is there. A schema theory of concepts, for example, presents the concept as the structured set of properties (i.e., the schema itself is the concept, and the schema is made of the properties in a certain structured set of relationships to each other). Property verification then, is a matter of somehow seeing whether the property is a part of the schema. The relation and reasoning approach and the more feature-based approaches to conceptual combination therefore treat property verification quite differently. We should note that in this discussion of the various models, we have grouped schema-based and property-based theories together because both approaches assume that properties are parts of concepts (that is, they both use a container metaphor of concepts). However, these two approaches differ in other aspects; for example, property-based theories were derived from propositional logic and schema-based theories were derived from predicate calculus (see Barsalou and Hale 1993). These differences, although important, are not relevant for the purposes of our current investigation.

Recently, a series of research projects have aimed at understanding property verification in combined concepts in order to evaluate competing theories of conceptual combination (Connolly et al. 2007; Gagné and Spalding 2011, 2014b; Jönsson and Hampton 2008, 2012; Hampton et al. 2011; Spalding and Gagné 2015). The most important result of these studies has been the modification effect: modification leads properties that are true of the head noun of the modifier-noun phrase to be less true of the modified concept (e.g., *baby ducks have webbed feet* is judged less true than *ducks have webbed feet*) while properties clearly false of the head are judged less false of the modified concept (e.g., *purple candles have teeth* is less false than *candles have teeth*), despite the fact that the properties were explicitly chosen to be unrelated to the modifier that was used. Indeed, the same effects arise even when

the modifier is a non-word (e.g., *chonk ducks have webbed feet*), and therefore cannot provide any conceptual content to act as the basis for the property verification judgment (e.g., Gagné and Spalding 2014b). Furthermore, the modification effect seems to be very robust over levels of typicality of properties of the head. That is, while features highly typical of the head are generally judged more true of the modified concept than features that are not highly typical of the head, typicality of the property for the head does not seem to provide much resistance to the modification effect. Even very highly typical properties of the head (e.g., the property *edible* for the concept *lamb*) show the modification effect (e.g., Hampton et al. 2011; Spalding and Gagné 2015). Finally, we should note that the modification effect appears to be quite robust over specific verification tasks, including ratings of likelihood of the truth of a property for a category, true/false decisions about the property's relation to the concept and the response times to make those decisions, and estimates of the percentages of category members for which the property is true.

(Gagné and Spalding 2011, 2014b; see also Spalding and Gagné 2015) argue that the modification effect is largely due to participants' expectation that subcategories will be both similar to and dissimilar from their categories, and that this suggests that property verification is more a result of reasoning than of inspecting the contents of the combined concept. Furthermore, they argue that the modification effect strongly suggests that properties are not transferred from constituents to the combined concept, but that instead the properties are predicated about the combined concept as a result of the reasoning that results from the demand for property verification.

Thus, we suggest that the processing of a combined concept initially involves constructing a structural description that denotes the way in which the two constituents are connected and determines the linguistic (morphological) structure in which the lexical representations of the constituents are mapped onto particular morphosyntactic roles. This assumption follows from linguistic theories that also propose a structure-based approach (e.g., Allen 1980; Levi 1978; Libben 1993). Thus, the phrase *cherry pie* begins as a modifier-head structure in which the *cherry* is mapped onto the modifier role and *pie* is mapped onto the head role. This linguistic information is involved in constructing a relation-based structure for the corresponding combined concept. The phrase *cherry pie* corresponds to *pie made with cherries*. Properties can then be inferred for this structure. This aspect of our proposal is consistent with previous data indicating that relational structures were available prior to properties. For example, Eliaser (1994) found that relation-based interpretations (e.g., whale boat = boat for hunting whales) were faster than property-based interpretations (e.g., whale boat = a big boat). Similarly, Gagné (2000) demonstrated that it took longer to judge that a property definition was acceptable than to judge that a relational definition was acceptable.

When inferring properties, participants use the meta-knowledge that XY's are a type of Y (see Bauer 1983) to determine that properties that are true of Y (i.e., of the category) are likely to be true of the subcategory, unless there is specific information that would suggest otherwise. However, participants also draw on meta-knowledge that subcategories are different in some way from the category.

Indeed, the purpose of using a combined concept rather than just the category is usually to refer to a subcategory that is distinct from other members of the category (e.g., Clark and Berman 1987; Downing 1977). This accounts for why previous research (e.g., Connolly et al. 2007, 2011; Gagné and Spalding 2014b; Hampton et al. 2011) has found that a property (e.g., has webbed feet) that is true of an unmodified concept (e.g., *duck*) also is seen as being true of a modified concept (e.g., *baby duck*) but less so. This explanation readily accounts for why a property (e.g., has teeth) that is false of an unmodified concept (e.g., *candle*) is judged as being more true of a modified concept (e.g., *purple candle*), as reported in a series of experiments reported by (Spalding and Gagné 2015). Both results (for the true and the false properties) can be explained by participants reasoning that the subcategory is distinct in some way from the category. Interestingly, the reasoning process appears to be more heavily dependent on meta-cognition and on pragmatic factors than on the specific content of the concepts involved. Consequently, content-free modifiers (i.e., unknown words) also produced a modification effect even though they had no content that could be used to assess specific similarity or dissimilarity with the category. Finally, Spalding and Gagné (2015) argue that if property verification for combined concepts involves such predication following reasoning and judgment, it is possible that such processes are also involved in all property judgments for single concepts, not just combined concepts. This, in turn, suggests a very different, more Aristotelian-Thomistic, approach to the relation between property and concept, rather than the kind of implicit “container” metaphor that is so common in current theories of concepts. In particular, in the A-T approach, concepts are not “made of” properties and they do not “contain” properties, and so one does not verify properties by “looking inside” the concept. We return to this point in the Discussion.

3 Current Experiment

In the current experiment, we investigate whether the modification effect is sensitive to the way in which a test property is initially predicated of the head. Spalding and Gagné (2015) used a procedure in which the likelihood of an unknown property was manipulated for a particular head concept, and participants were then asked what percentage of the head concept’s category members had the property or they were asked what percentage of the members of the modified concept had the property. For example, the participants might be told, “Almost all birds have sesamoid bones. What percent of birds/baby birds have sesamoid bones?” Using this procedure, Spalding and Gagné found robust modification effects: The percentage estimates for properties presented as true of almost all category members were higher for the unmodified than for the modified concepts. Spalding and Gagné also found a robust inverse modification effect: When the property was presented as true of almost no member of the category, the percentage estimate was lower for the unmodified than the modified concept. Spalding and Gagné’s manipulation was

about the likelihood of the property for the unmodified head. However, it is unclear whether properties predicated in ways that are more normative, rather than just about the likelihood of the property, would show the same modification effect (Prasada and Dillingham 2009, for example, have shown that simple statistical likelihood does not account for all aspects of the relation of properties to concepts).

In particular, we were interested in whether properties predicated with varying degrees of normative force would show variation in the size of the modification effect. Investigating this point will allow us to determine the extent to which the inclusion of properties is affected by general meta-knowledge about the parent category (e.g., about the category *bird* when reasoning about a subcategory of birds). We chose three ways of predicating the properties. In the Bare Generic condition, the test property was predicated generically with no modification: *birds have property X*. In the Normative condition, the test property was predicated in a normative way: *normal birds have property X*. In the Typical condition, the test property was predicated in a less normative, but frequentist way: *typical birds have property X*. Thus, the Bare Generic condition should have the most normative force, and should thus result in the highest percentage estimates. The Normative condition relaxes the implied “every” from the generic condition, but maintains some degree of normative force, implying that a bird lacking property X might be abnormal. Thus, one might expect that properties predicated in this condition would have lower percentage estimates than those predicated in the Bare Generic condition. The Typical condition further relaxes the force of the predication, such that a bird lacking property X might be atypical, but not abnormal. Thus, one might expect that the properties here would have the lowest percentage estimates of all the predication conditions. The same kinds of predictions could be made for false features (birds do not have property X, normal birds do not have property X, typical birds do not have property X), but with the reversed order (i.e., properties presented as generically false would have the lowest percentage estimation, then those presented as normative, then typical). In other words, the nature of the predication could be seen as directly affecting the strength of the property for the head concept, but without depending on the participants’ specific conceptual knowledge.

The critical interest for the current project, however, is whether this manipulation might affect the modification effect itself. One might make predictions for the size of the modification effect that parallel the predictions for the level of verification. That is, one might think that there should be little or no modification effect for the Bare Generic predication, a somewhat larger modification effect for the Normative condition, and the largest modification effect in the Typical condition. Again, similar, but reversed effects might be predicted for the false properties, across the predication conditions. On the other hand, given the previous work showing, for example, that even categorical properties (e.g., *lambs are mammals*; Hampton et al. 2011) show a robust modification effect, and in which the major factor affecting the modification effect was the expectation of contrast between the category and the subcategory rather than the specific content of the categories (e.g., Spalding and Gagné 2015), one might expect robust and equivalent modification effects across all three predication conditions.

3.1 Method

Materials and design. 84 head nouns were drawn from Gagné and Spalding (2011). Unmodified (*bottles*) and modified nouns (*safety bottles*) were constructed. An unknown predicate (i.e., a potential property of the head which the participant was not expected to know) was associated with each noun. Three levels of normative force were chosen to manipulate the type of predication for the unmodified noun: Bare Generic (e.g., *bottles are cooled in annealing ovens*), Normative (e.g., *normal bottles are cooled in annealing ovens*), Typical (e.g., *typical bottles are cooled in annealing ovens*). A second set of each type of predication was created to present the same unknown property as false of the unmodified noun (e.g., *bottles are not cooled in annealing ovens*; *normal bottles are not cooled in annealing ovens*; *typical bottles are not cooled in annealing ovens*).

The test items were questions asking what percentage of the modified or unmodified nouns had the property (e.g., What percentage of bottles/safety bottles are cooled in annealing ovens?). Predication type was crossed with Property (True/False) and Modification (Unmodified/Modified) to create a $3 \times 2 \times 2$ design.

Participants responded using a scale with a slider that moved from 0 to 100%. These percentage estimates were the dependent variable in the experiment. The materials were counterbalanced such that each level of predication was tested equally often with each modified and unmodified noun. Each participant saw each head noun on only one trial during the experiment. Thus, each person saw 84 trials.

Participants. One hundred forty-four first-year psychology students participated for partial course credit. All were native English-speakers.

Procedure. The instructions included information about the unknown predicates: “The goal of this research is to understand the conceptual system. In everyday life, we are frequently called upon to make judgments of the likelihood of something being true on the basis of limited information. Consider the following situation: If you were told that mammals require biotin for hemoglobin synthesis, how likely do you think it would be that dogs require biotin for hemoglobin synthesis? Because some concepts are described by one word (e.g., dogs), whereas other concepts are described by phrases (e.g., hunting dogs), in this study you will be asked to make judgments involving both types of concepts.”

On each trial, the participants were given a statement in which an unknown predicate was presented with an unmodified noun and participants were told to assume that the statement was definitely true. They were then asked what percentage of either the unmodified or modified noun would have the property (e.g., What percentage of bottles/safety bottles are cooled in an annealing oven?).

Table 1 Mean (SE) estimations of the percent of category members having the test property by property type, predication type, and modification

Condition	Bare generic	Normative	Typical
	True: percent estimation (SE)		
Unmodified	96.7 (0.41)	88.2 (0.52)	88.3 (0.47)
Modified	92.6 (0.61)	83.3 (0.82)	82.7 (0.83)
	False: percent estimation (SE)		
Unmodified	2.1 (0.34)	13.1 (0.66)	12.3 (0.61)
Modified	4.7 (0.52)	16.3 (0.81)	16.5 (0.79)

3.2 Results

The descriptive statistics for the percentage estimates are presented in Table 1. The data were analyzed using Linear Mixed Effects models, as these allow participants and items to be simultaneously treated as random effects (thus avoiding the “language as fixed effect fallacy”, Clark 1973), among other advantages (for discussion, see e.g., Pinheiro and Bates 2000; Quené and van den Bergh 2004, 2008). Predication condition (Bare Generic, Normative, Typical), Property (True, False), and Modification (Unmodified, Modified) were entered as fixed factors. Interactions involving Predication were tested using Chi-square tests of model fits, while other effects involving only factor comparisons with two values are reported as z-scores, based on the slopes of the predictors in the models. The analyses were conducted using the *mixed* function in Stata 13.

Most critically, there is no three-way interaction among Predication, Property, and Modification, $Chi^2(2) = 3.11$, $p = 0.21$. Thus, there is no evidence that the modification effect differs across the levels of Predication and Property.

Because the modification effect is (as predicted) in opposite directions for the true and false properties, it is easiest to understand the data by separating the data on this factor. When considering only the true properties, there is no interaction between Predication and Modification, $Chi^2(2) = 1.99$, $p = 0.37$, but there are strong effects of both Modification ($z = -5.27$, $p < 0.001$) and Predication, with the Bare Generic condition differing from both the Normative ($z = -11.0$, $p < 0.001$) and the Typical ($z = -11.1$, $p < 0.001$) conditions. There is no evidence that the modification effect is sensitive to Predication type.

When considering only the false properties, there is again no interaction between Predication and Modification, $Chi^2(2) = 1.75$, $p = 0.42$, but there are strong effects of both Modification ($z = 3.08$, $p < 0.002$) and Predication, with the Bare Generic condition differing from both the Normative ($z = 12.0$, $p < 0.001$) and the Typical ($z = 13.0$, $p < 0.001$) conditions. As was the case for the true properties, there is no evidence that the modification effect is sensitive to Predication type.

Finally, it is perhaps important to note specifically that there are robust modification effects in the Bare Generic condition, for both true ($z = -6.49$, $p < 0.001$) and false ($z = 4.44$, $p < 0.001$) properties. Thus, although the Bare Generic condition has numerically slightly smaller modification effects, there is no statistical

evidence that those effects differ from the modification effects in either the Normative or Typical condition, and there is evidence for a very robust modification effect, even in the Bare Generic condition.

4 Discussion

The results indicate that the type of predication does matter for property verification, with the Bare Generic form of predication having the highest percentage estimates for the true properties and the lowest for the false properties. This suggests that participants were treating these statements as being the most canonical in that the statements were conveying information that most members of that category share. Indeed, the Bare Generic condition percentage estimates in the current experiment are very similar to those obtained by (Spalding and Gagné 2015, Experiment 2) when they told participants that “almost all” members of the head noun category had the property and then asked participants to estimate percentages. The task used in Spalding and Gagné was identical to the one in the current experiment, except that the likelihood of the property being true of the category was manipulated by using Almost all, Some, and Almost No, suggesting that the Bare Generic predication encourages a reading that is similar to “almost all”, at least in the context of the other predication types. However, the Normative and Typical types of predication did not differ from each other, for either true or false properties. This suggests that participants were treating both the “normal” and “typical” statements as linguistic hedges, rather than as applying differing normative force (as was the case for the Bare Generic form). Indeed, it might be the case that these conditions simply reminded the participants that there are many abnormal and atypical members of most categories, thus lowering the overall likelihood of the property. Thus, it appears that information provided by the type of predication (Bare Generic, Normative, and Typical) is used to provide a starting point for what percentage of category members are expected to have a given property.

Although the type of predication affected the percentage estimates strongly, it had no impact on the modification effect itself (as indicated by the lack of interaction). Thus, as with high levels of typicality or centrality (Hampton et al. 2011; Spalding and Gagné 2015), presenting the property as generically true (or false) of the unmodified head does not provide any resistance to the modification effect, relative to presenting it as normally or typically true (or false). For example, in Hampton et al. (2011), the modification effect was the same for statements with mutable properties (e.g., lambs are white versus baby lambs are white) and immutable (or central) properties (e.g., lambs are warm blooded versus baby lambs are warm blooded). Interestingly, in the current data, despite the fact that the normative and typical framings of the predications resulted in much lower (higher for false properties) percentage estimates, this did not seem to make those conditions any more susceptible to the modification effect than the bare generic condition, as one might expect if those conditions simply made the participants think that

the categories were more variable with respect to those properties. This result (as well as the lack of interaction in terms of mutable versus immutable properties observed in previous studies) is consistent with the view that the modification effect is primarily due to participants' expectations of partial contrast between categories and subcategories (Gagné and Spalding 2011, 2014b; Spalding and Gagné 2015). That is, because those true properties are still estimated at about 85% for the unmodified head (and false properties at about 15%), they drive an expectation of contrast when the noun is modified, but no more nor less of an expectation of contrast than the bare generic condition. To further illustrate, it appears that inferential processing is drawing more heavily on meta-knowledge about the function of modification rather than on content-specific knowledge about the particular properties and concepts, and consequently, for true properties, participants reduce the estimate for modified concepts relative to unmodified concepts due to the expectation that subcategories differ in some way from categories. An estimate about any statement that has the form XY has property P is reduced (for true properties) or increased (for false properties) relative to statements of the form Y has property P. As discussed in the previous paragraph, the starting estimate (i.e., the estimate for the unmodified concept) can differ depending on the form of the prediction, but the size of the increase or reduction is independent of this starting estimate.

To give a concrete example, when participants are told that *bottles are cooled in annealing ovens*, they assume that most members of the *bottle* category have that property. They draw on this information and the meta-knowledge that an XY tends to have the same properties as a Y when estimating what percentage of *safety bottles* have that property to initially consider a relatively high percentage. This estimate is then lowered because participants also use the meta-knowledge that subcategory members are distinct, in some way, from the category. Interestingly, even though there is nothing inherent in the content of the modifier to conflict with that property (just as there is nothing in the content of *baby* that would prevent a *baby lamb* from being warm-blooded), the bias towards viewing categories and subcategories as partially contrastive led to the lowering of the estimate of a property that was previously said to be true of the category. The application of this same bias explains why participants *increase* the estimate for a property that is said to be false of the category; the property that is not true of most members of the category is seen as the basis of the difference between the subcategory and category and, thus, subcategory members are thought to be more likely to hold this property.

One might expect, on this view, that there should be an interaction of modification with predication type, such that the modification effect would be larger as the likelihood of the property for the head goes down (Hampton, personal communication). That is, in the limiting case, if 100% of the category has the property, then 100% of the subcategory should have it, and as one moves away from the 100% mark for the category as a whole, then there is more "room" for the subcategory to contrast with the category (while allowing the percentages to still work out from a mathematical perspective). Two considerations argue against this, however. First, it is highly unlikely that the participants are calibrating their estimates this finely;

instead they seem to simply apply a more or less equivalent level of adjustment of likelihood in all cases. Also, of course, it is known that people are not particularly good about making sure that probabilities or percentages “add up” properly (see, e.g., Jönsson and Hampton 2006). Second, even if they were calibrating their estimates this finely, there is another effect that would seem to be a countervailing force: (Spalding and Gagné 2015) found that when the property was not true enough of the head category, then there was no modification effect. In short, the expectation of contrast depends on there being an initial expectation that the property either is or is not true of the unmodified concept. When participants were told that some X’s have property P, there was no modification effect (nor was there an inverse modification effect). Generally, of course, the existence of both the modification effect and its inverse assumes that at some level of likelihood, there is a reversal of contrast expectation, and hence there are probably some levels of likelihood where there is no effect. In other words, as likelihood of a property decreases towards 50% or so, the modification effect goes away, and then re-emerges as an inverse modification effect as the likelihood of the property moves further toward 0%. Thus, even if participants were trying diligently to make their percentages add up properly, we should be unlikely to see the kind of interaction suggested above, as this change to expected contrast works in opposition to the interaction described in the earlier part of this paragraph.

4.1 The Aristotelian-Thomistic Approach

As suggested by Gagné and Spalding (2011,2014b) and Spalding and Gagné (2015), the effect of expected contrast on the modification effect suggests a very different kind of property verification process than that envisaged in other research on this topic (Connolly et al. 2007; Jönsson and Hampton 2008, 2012; Hampton et al. 2011). In particular, the results suggest that conceptual combination does not consist of the merging and evaluation of properties of the constituent concepts as the mechanism for creating new concepts. That is, the creation of the concept *farm dog* does not appear to consist of accessing properties of *farm* and of *dog* and then building the concept *farm dog* from these properties. If this were the case, then it is difficult to explain the pattern of results observed in the current experiment and in previously published work on the same topic (such as Spalding and Gagné 2015, which also reports an inverse modification effect for false properties, such as *(purple) candles have teeth*). In particular, it is more difficult to understand why properties that were explicitly said to be not true of the category would be added to the subcategory. Furthermore, it strongly suggests that the process of conceptual combination is not a process of mechanically transferring properties from constituent concepts to the newly combined concept, but instead reflects reasoning at the point of decision (for other arguments against such a property transference mechanism, see Connolly et al. 2007; Fodor 1994, 1998, 2001; Fodor and Lepore 1993, 1996). Indeed, these results are difficult to explain using feature-based

theories of conceptual combination. This is particularly true for the false properties, which would have no reason to arise in the combined concept's representation under any of the theories of conceptual combination. The modification effect also suggests that property verification is not a matter of looking into the concept to see if the property is there. Now, as Spalding and Gagné point out, if one does not do property verification by inspecting the contents of the combined concept, perhaps one also does not make property decisions about any concept by inspecting the contents of the concept. How, then, does one make such decisions?

We believe that the current results, and the previous results of research on the modification effect, point to the need for an alternative framework, based on the incorporation of reasoning processes and meta-knowledge about the role of modification in human communication, that could be used in forming and testing psychological theories of concepts and conceptual combination, and in particular theories of the relationship between properties and concepts. Surprisingly, there is relatively little modern research specifically into the relationship between properties and concepts (see, e.g., Laurence and Margolis 1999). Historically, however, the relation between properties and concepts has been investigated in great detail in philosophy, particularly in the work of Aristotle and Thomas Aquinas. (Spalding and Gagné 2013) have argued that the Aristotelian-Thomistic (A-T) view of concepts might have much to tell us about concepts (see also Spalding and Gagné 2015). They showed that the A-T view is not the same as the modern "classical" view, which was (correctly) ruled out by concepts research in the 1970s and 1980s, and discussed several aspects of modern concepts research that fit surprisingly well with the A-T understanding of concepts. Therefore, we end this chapter with a description of properties and concepts in the A-T view, and the suggestion that taking this philosophical view of concepts seriously could have real benefits for research in the psychology of concepts, and particularly for our understanding of the relationship between properties and concepts.

To begin, it is important to realize that thinkers in the A-T tradition have thought deeply about properties and concepts and about how one reasons about things in the world, and have made many distinctions at levels of granularity that are far finer than those in modern concepts research. Also, it should be noted that making judgments about properties is not limited, in the A-T view, to philosophy of concepts, but also explicitly involves the philosophical domains of logic and reasoning. The discussion here is necessarily brief and introductory, but there is much serious discussion of A-T thinking on properties (see, e.g., Reynolds 2001; Wipfel 2000, Part 2), concepts (for an overview, see, e.g., Mercier 1950a) and in the related judgments people make (see, e.g., Mercier 1950b). Indeed, some of the terminology in common modern use is taken from this tradition, though having lost much of its technical meaning in the transition. *Property* is perhaps the most obvious case. *Property* in modern usage is a part of the concept, and as such is a representation of any characteristic believed to be associated with members of the category denoted by the concept. *Property* in the A-T tradition is quite different. First, *property* is a shorthand for *proper accident*, and a proper accident always inheres in things in the world, not in concepts. *Accidents* contrast with *substances*. Substances, to a first

approximation, are things that exist in themselves, while accidents are things that exist in or ‘inhere in’ substances. Thus, neither substances nor accidents (including proper accidents) primarily exist in representations because accidents inhere in substances and substances are found in the real world. Furthermore, proper accidents contrast with other kinds of accidents, in that proper accidents “flow from” the essence of the substance in which they inhere, while other kinds of accidents do not flow from the essence of the substance. For example, having two eyes flows from the essence of being human, but having brown hair does not. Note that being necessary, in this sense, does not mean that every existing human must have two eyes, or that a human who loses an eye becomes non-human, but rather that it is in the nature of the human to have two eyes, and that a failure to have two eyes is a deprivation of the normal property flowing from the essence. Such deprivations are well-known and well-explained by the A-T view (see Spalding and Gagné 2013, for more discussion and comparison to the so-called classical view of concepts). Thus, properties, in the A-T view, are not parts of concepts. Instead, they are characteristics of things in the world, though one may have a concept of a given property, of course. Properties can be predicated of primary substances (e.g., *wings* can be predicated of individual birds in the world), or of secondary substances (i.e., *wings* might also be predicated of a universal such as *birds*). Also, such predications can be remembered and become habitual, because we have understood how *wings* and *birds* relate. Hence, one need not consciously think through the relationship of *wings* and *birds* anew every time. Importantly, though, it is the act of predicating *wings* of *birds* that becomes habitual. Thus, it is not the case that *wings* must become part of the concept *bird*, even when the predication happens habitually (i.e., in the A-T view, it is not necessary to turn to a conceptual container approach to account for the predication).

A second important characteristic of A-T thinking about concepts and properties is that, unlike most thinking about concepts since at least Locke, the concept is not the thing that the mind considers when it thinks about something. Rather, the thing is what the mind considers, while the concept is simply what allows the mind to think about the thing. One thinks about the external object via the concept, rather than thinking terminating in or being directed toward the concept itself. Of course, the mind can reflexively think about a concept—I can consider my own concept of *bird* and reflect upon what it is to have a concept at all, but that is not what I do when I think about birds; note that this distinction is rather hard to keep clear under a Lockean approach, where to think about birds just is to activate the concept *bird*, and thinking about the concept *bird* is also to activate the concept *bird*. Relatedly, thinking, in the A-T tradition, is sometimes said to involve representations, but this is not quite correct, in modern terms. Instead, thinking in the A-T tradition involves *re-presentations* of the thing to the mind, of the mind *re-presenting* the thing internally. Thus, *re-presentation* (in the verb sense) is primary in the A-T view, whereas a *representation* (in the noun sense) is primary in modern theories. In the A-T view, representation of a thing is an event, an action undertaken by a cognizer through which a thing is known, not typically a thing known in itself. Although this distinction may seem picayune, it is important to note that the early modern (and

continuing) difficulties with radical skepticism and the “problem of the bridge” derive from the claim that the ideas are what the mind grasps—If the mind grasps ideas, how do we bridge from the ideas to the world? What becomes a fundamental problem for many modern accounts of knowledge is simply not an issue on the A-T view of knowledge, since in this account the mind grasps the thing through the concept (or, stated differently, the thing is re-presented through the concept) rather than grasping the concept itself. For Aristotle and, especially, Thomas Aquinas, the mind mirrors or re-presents the external world in a basically reliable way (Feser 2009, pp. 145–149, provides a clear discussion of this point; for a more detailed treatment, see Perler 2001).

A third important characteristic of the A-T view of concepts and properties is that all thinking involves the use of sensory information. In the A-T view, any judgment about a thing, for example, involves not only the concept of the thing, but the *phantasm*, the sensory information associated with the actual things that the concept is “of”. There is, following a famous dictum of Aristotle, no thinking without a phantasm. This necessary involvement of the phantasm is important because it means that there is a constant, fundamental, and completely thorough-going recourse to the senses for any thought whatsoever. The A-T view has a long history of carefully considering how exactly sensory and intellectual information relate and interact in thinking. This history could be very helpful in understanding some of the recent work on embodied concepts and perceptual symbol systems (e.g., Barsalou 1999). Indeed, notions such as the re-enactment of conceptual content via simulation (e.g., Barsalou 2009, p. 1282) accord extremely well with the A-T notion of re-presentation (verb) as opposed to representation (noun). Similarly, the way that (Wu and Barsalou 2009) describe the process of generating features from simulations is very similar to how the A-T view would describe the process of predicating properties via the phantasms (which are also multi-modal and sensory-based). Furthermore, the A-T approach incorporates the imagination, a power of manipulating phantasms, such that sensory information can be retrieved, manipulated, and re-combined in novel ways, as well as common sense (here a technical A-T term), which combines sensory information from multiple sensory modalities. Finally, conceptual thinking in the A-T view involves the intellect, which is the power called upon in formal reasoning, abstract concept formation, mathematics, and other such tasks, as well as a power that can direct the imagination and access the phantasms. Thus, when a person thinks about concepts and properties, all of these various powers come into play.

A fourth important characteristic of the A-T view is that properties (or features) are not more fundamental than substances. Rather, substances are more fundamental than properties, because we say or ‘predicate’ properties of substances and not the other way round. For example, we predicate ‘rational’ of ‘Socrates’ and not ‘Socrates’ of ‘rational’. We would never claim that ‘Rational is Socrates,’ though the obverse clearly obtains, namely ‘Socrates is rational.’ The property (rationality) is predicated of the substance (Socrates). Moreover, we never predicate one substance of another. For instance, we cannot predicate Socrates of Plato in the following judgment: Plato is Socrates. Instead, we predicate various accidents of

substances and these accidents or features are seen to inhere in the substance. This contrasts markedly from most modern views, in which concepts consist of properties. That is, the concept is made of or just is the set of properties. Thus, the properties are more basic than the concept, which has several interesting consequences. One is that if this general view were true, one would expect that categorization would be a matter of identifying the properties of a thing and then using those properties to determine whether this thing is in the category. Indeed, as a general description, this is basically how categorization is understood in most modern theories. However, this implies that the properties should be identified prior to the category. Thus, before one grasps that the thing they are looking at is a bird, they should already have grasped that they are seeing a beak and feathers and wings, etc. Note that it is not that they are seeing visual information that can be interpreted in these ways, but literally that they are grasping those conceptual features. A second implication of this approach is that there should be a clear and simple linkage from the extraction of simple visual features to simple conceptual features to the concept. In short, we should have no discontinuity between research on object recognition and research on concepts and categorization. Yet, research on object recognition rarely uses conceptual features, while research on concepts and categorization assumes conceptual features and rarely even posits the exact connection to the simple visual (or, for that matter, other sensory) features.

Finally, it is important to understand that the A-T view of concepts and properties is embedded within a very complicated and intricate understanding of not only the mind, but of the physical world. Thus, for example, the A-T way of thinking about essences, properties, and concepts derives not from what is convenient for philosophy of mind, but from prior metaphysical commitments about the physical world and how the mind is seen to function (fairly reliably) as a part of that world. That is to say, the A-T approach to accounting for physical change through terms like actuality/potentiality and form/matter serves as the basis for A-T cognitive theory, and this cognitive theory is, in turn, a further specification of those basic metaphysical commitments. An immediate and obvious benefit of this approach, say advocates of A-T cognitive theory, is that it is completely consistent across differing levels of cognitive and metaphysical explanation, from how intellectual concept formation occurs, to other sorts of cognition, to an explanation of the constitution of physical things (including humans), and even to the processes of change as such (e.g., Feser 2014) because it is all part of a general and consistent view of the world.

What is surprising about this is how well the philosophy of mind inherent in the A-T approach fits modern evidence about the mind. (Spalding and Gagné 2013), for example, discuss how the A-T approach fits well with a wide variety of evidence drawn from modern psychological research on concepts, including evidence for prototypes and exemplars, as well as work on generics and essentialism, the relation between human and non-human animals' abilities to deal with categories, and other recent research topics in concepts. We should also note here that (Prasada and Dillingham's 2009) notion of K-properties (i.e., properties that are indicative of a kind of thing) is explicitly Aristotelian, though they do not explicitly make use of

the rest of the A-T approach to cognition. Pustejovsky (1995) incorporates a version of the four Aristotelian causes in his “qualia” structure for lexical items. Nor should we think of the A-T view as being relevant only to the study of concepts or language. Feser (2014) provides a broader discussion of the A-T approach as an appropriate underpinning for science in general, both in the sense that its view of human cognition provides a proper underpinning for the human activity of doing science and in the sense that its metaphysical commitments provide a proper underpinning for science considered as a set of propositions about the world. Indeed, the A-T view may well provide at least as good a philosophy of mind as we can find in modern and contemporary approaches (see e.g., Madden 2013; see also Spalding et al. 2014). At the very least, the consistency between cognitive theory and metaphysical commitments in the A-T view presents a picture that contrasts sharply with modern psychology and philosophy of mind, where there seem to be extraordinarily deep and intractable problems in consistently linking down to metaphysics (Madden 2013, for example, discusses many problems with the relations between modern philosophy of mind and modern metaphysical approaches; see also Feser 2006, 2008, 2009).

4.2 Applying the A-T View to the Modification Effect

The applicability of the A-T view to cognitive science is seen acutely in what we have been discussing throughout this article: property verification judgments and subsequent modifications made to those judgments. To show this applicability, it is perhaps best to begin with a general explanation of how a property verification judgment is made (in its strictest sense) on the A-T view (i.e., through what Aristotle and Aquinas call *episteme* or *scientia*, respectively), and then present how the same approach might account for the sort of property verification results encountered in our experiment.

On the A-T view, property verification in its strictest sense occurs in virtue of understanding how a given property is explanatorily useful in knowing the nature of the thing in which the property is found. In other words, in the A-T view, to do property verification is to verify that a characteristic is, in fact, a property of the thing (i.e., that it is a characteristic that flows from the essence of the thing). To put the matter concretely, we can ask, ‘How does one decide whether birds have wings?’ but with the understanding that we are actually asking how one decides whether *wings* is a proper accident (flowing from the essence) of *bird*. On the A-T view, property verification in this strict sense will occur when we understand why the property regularly occurs in the kind of thing of which we are predicating the property. Note that in the A-T view, much of what we “know” about the world, even though a reliable guide for action, etc., is not actually known in the strict sense of *episteme* or *scientia*. We return to this point below.

Now, as we have pointed out above, judgments are not made by looking in the concept to see if there is a property there, but rather through a judgment about the

thing to be known (thus, the distinction between a concept as what the mind grasps vs. that by which the mind grasps the thing to be known, is quite critical to a proper understanding of how property judgments are made). In the A-T view, our knowledge begins in a confused and indeterminate manner and grows more specific as we encounter a thing regularly (if we choose to try to understand). As we regularly encounter birds, we may notice wings, beaks, feathers, etc., and notice that other animals for which we have refined concepts do not have these characteristics. We do not yet know what these characteristics are, nor what they are for, but they may make the object of inquiry stand out with respect to other animals. Perhaps after further inquiry we come to see what these features are for and notice that many or even all birds that we have encountered have these features. Once we have this knowledge (i.e., what these features are for and a wide, perhaps, universal distribution), wings, beaks, and feathers will become part of the causal story of birds, and we can understand why birds have wings (even, with much inquiry, flightless birds). At this point, we have recognized and predicated a proper accident, a property that we think falls to all birds because we see that the nature of ‘bird’ is at least partially, but essentially, explained by the property ‘wing’. We thus engage in a process of developing more precise knowledge as we become more familiar with a given object. Ultimately, in the A-T view, this process can (but does not always) result in a definition, in which the thing can be defined, in terms of its nature, via its genus and specific difference (such that the human being can be defined as a type of animal, specifically a rational animal; note that genus and species are used here as technical terms, but not in their modern biological technical senses).

How, then, would the A-T view explain the sort of property verification that we tested in our experiment, where we are clearly not asking for property verification in the strictest A-T sense? That is, how does one, on the A-T view, account for the judgment that ‘birds have sesamoid bones,’ and, more pressingly, then account for a modification effect when test subjects were told to apply ‘sesamoid bones’ to a newly combined concept, like ‘baby bird’? These examples clearly differ from the property verification in the strict sense described above, and for at least two reasons: first, our test subjects do not know the meaning of the predicate (in this case, ‘sesamoid bones’), whereas by the end of investigation the inquirer, by definition, knows what the predicate means in what is described above. Second, A-T property verification in the strict sense is meant to result in a ‘categorical’ judgment, like ‘*All* birds have wings,’ which applies universally to all things in the relevant category (note, however, our comments above that terms such as “universal” or “necessary” do not actually mean that every existing example will have the properties—there is still the issue of deprivation of such properties—rather it means that such properties are seen to flow from the essence). By contrast, in our experiment, our test subjects are operating on distinctions of ‘generic’, ‘normal’, and ‘typical’, none of which claims that the property universally applies. With those differences in mind, we think that the A-T approach to cognitive theory is consistent with our empirical results regarding reasoning about combined concepts.

An essential part of this story is to account for how and why reason functions as it does when applied to combined concepts, which plays to the strengths of the A-T

view. While the A-T view of cognition has a robust (proposed) philosophical description of the mechanics of cognition (see, e.g., Brennan 1941; Kosteletzky 2014), from how sensory data is initially perceived in a phantasm to an intellectual judgment about the thing, it does not rest solely on those mechanics. Instead, there are different sorts of knowledge and different kinds of inquiry that supervene on the process that results in the conceptualization of the thing. The ‘highest’ form of knowledge (‘highest’ because it is most certain) is simply ‘understanding’. In this form of knowledge, one’s assent is forced, as it were, and there is no process of reasoning required for assent. A classic example is ‘A whole cannot be smaller than one of its parts.’ If I have understood the terms at play, I cannot force myself to assent otherwise, since it is impossible to conceive of a whole being smaller than one of its parts, and, importantly, the truth of the statement is immediately apparent; reasoning and inquiry do not occur. Below this sort of knowledge is ‘science’, which is the so-called strict sense of property verification discussed above. ‘Science’ deals in necessary knowledge, with conclusions that cannot be otherwise than they are, but reasoning is clearly employed because ‘science’ follows the pattern of a demonstrative syllogism. For example, if I know that all humans are rational and that all Greeks are human, I know that all Greeks are rational.

Below this sort of inquiry is ‘dialectics’, which also makes claims that are meant to be universal, but the premises are only taken as likely true, because the premises are acquired or accepted from “the received opinion of the wise and the many” (Aristotle, *Topics* I). This basic aspect of dialectics dovetails nicely with how our research participants received their knowledge that ‘bottles are cooled in annealing ovens’ or that ‘birds have sesamoid bones.’ We told them, and they took this received opinion as true. Nevertheless, on the A-T view they neither ‘understand’ the proposition nor have an account of the ‘reason why’ the proposition is true. It is, at best, ‘probable knowledge’ (i.e., what Aristotle claims is the product of dialectics). Yet, nearly all our participants assented that these properties were true of the head, as can be expected given that we (“the wise and the many”, at least in the context of our experiment!) asserted the truth of these propositions.

As we modified the concept, a modification effect comes into play, and fewer subjects assent to the proposition involving the modified concept. This can be accounted for by way of the relationship between reasoning and assenting to a proposition, because assent, for Thomas Aquinas (he discusses issues of assent directly, whereas Aristotle does not treat it as precisely), occurs more firmly and immediately if the premises themselves are more certainly known. If the premises are not firmly known, more reasoning is required to get to the point of assent (see Thomas’ Disputed Questions on Truth, q. 14, art.1). Even then we are going to hold less firmly the proposition to which we have assented.

On this account of how reasoning and assent function together, it is hardly surprising that fewer people will assent to a proposition with a modified concept, especially if they took the truth of the original proposition on trust. There are two reasons for this. First, of course, we clearly have only probable reasoning in moving from the truth of the property for the head to the truth of the property for the combined concept, and this, according to Thomas, should result in lower confidence

in the outcome of that reasoning. This aspect of reasoning to assent would apply to any proposition that involved probable reasoning (i.e., the strength of the conclusion is going to be lower than the strength of the premise whenever the reasoning involved is only probable, unless direct knowledge of the conclusion, or other knowledge, can intervene).

Second, and more specific to the modification effect per se, the structure of the Thomistic understanding of category and subcategory should give rise to the kind of expectation of similarity and contrast suggested by Gagné and Spalding (2011, 2014b) and Spalding and Gagne (2015). That is, as noted above, in the Thomistic view a subcategory is distinguished from a category via a specific difference (i.e., something that differentiates this species from its genus, as in the human being defined as a rational animal). Hence, one should expect that there is some difference (and, likely, an essential difference) when a subcategory is being posited. Of course, a person can posit a purely accidental or even meaningless subcategory, but that should be a relatively rare event and should not much affect one's expectation that a posited subcategory would differ from the category in meaningful way. Even in the case of a purely accidental subcategory (in A-T terms, i.e., a subcategory that differs primarily in non-proper accidents), there will often be reasons for the non-proper accidental difference that will be related to the essence of the category. Consider, for example, the subcategory of birds that live in Edmonton, Alberta, Canada. Although there is no particular essence that relates to this rather artificial and accidental subcategory, there are differences between the birds that can live in this location and those that cannot, and these differences would be very likely related to the essences of those various types of birds. Hence, one might expect that this group may differ from birds in general in essential ways, even though this grouping is accidental. But, in many, if not most, cases of known subcategories there would be an explicitly essential difference—the essence of a subcategory just is not, generally, the same essence as the category. Furthermore, because multiple accidents (both proper and otherwise) will either flow from, or co-occur with, a difference in essence, there is good reason for the participant to expect contrast on various characteristics, but also to expect that overall, there should still be many similarities (because much that is essential, as well as proper and accidental, is in both the category and subcategory). In a nutshell, this is how (Gagné and Spalding 2011, 2014b; Spalding and Gagné 2015) account for the modification effect. It is not inherently about modification or conceptual combination, but instead is about the expected relationship between the category and the subcategory. That expected relationship of similarity and contrast enters into the reasoning process that the person goes through in evaluating the likelihood of the truth of the predication of the property to the subcategory.

5 Conclusion

The current experimental results are consistent with a view of the modification effect in which the primary driver is the expectation of similarity and contrast between category and subcategory, as suggested by (Gagné and Spalding 2011, 2014b; Spalding and Gagné 2015). In particular, and contrary to what would be expected given feature-based theories of conceptual combination, the modification effect seems to be insensitive to the normative force with which a property is predicated of the head noun concept, though the overall level of attribution of the property is quite sensitive to the force of the predication. This view, in turn, is consistent with an approach to concepts deriving from the Aristotelian-Thomistic tradition (Spalding and Gagné 2013). We argue that taking the A-T view seriously has much to offer current research on concepts and might provide a fruitful basis for the development of a psychological theory of conceptual combination and of concept processing, in general.

References

- Allen, M. R. (1980). Semantic and phonological consequences of boundaries: A morphological analysis of compounds. In M. Aronoff & M. L. Kean (Eds.), *Juncture* (pp. 9–27). Saratoga, CA: Anma Libri.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society*, 364, 1281–1289.
- Barsalou, L. W., & Hale, C. R. (1993). Components of conceptual representation: From feature lists to recursive frames. In I. van Mechelen, J. Hampton, R. Michalski, & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 97–144). San Diego, CA: Academic Press.
- Bauer, L. (1983). *English word-formation*. Cambridge, UK: Cambridge University Press.
- Brennan, R. E. (1941). *Thomistic psychology: A philosophic analysis of the nature of man*. New York: Macmillan.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12(4), 335–359.
- Clark, E. V., & Berman, R. A. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language*, 14, 547–567.
- Connolly, A. C., Fodor, J. A., Gleitman, L. R., & Gleitman, H. (2007). Why stereotypes don't even make good defaults. *Cognition*, 103, 1–22.
- Downing, P. (1977). On the creation and use of English compound nouns. *Language*, 53(4), 810–842.
- Eliaser, N. M. (1994). Distinct processes in conceptual combination. *Unpublished master's thesis, Northwestern Univ., Illinois*.
- Feser, E. (2006). *Philosophy of mind: A beginner's guide*. Oxford: Oneworld.
- Feser, E. (2008). *Last superstition*. South Bend, IN: St. Augustine's Press.
- Feser, E. (2009). *Aquinas: A beginner's guide*. Oxford: Oneworld.
- Feser, E. (2014). *Scholastic metaphysics: A contemporary introduction*. (Editiones Scholasticae: Vol. 39). Heusenstamm: Editiones Scholasticae.
- Fodor, J. (1994). Concepts: A potboiler. *Cognition*, 50, 95–113.

- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. New York, NY: Oxford University Press.
- Fodor, J. A. (2001). Language, thought and compositionality. *Mind and Language*, 16, 1–15.
- Fodor, J. A., & Lepore, E. (1993). Why meaning (probably) isn't conceptual role. *Philosophical Issues*, 3, 15–35.
- Fodor, J. A., & Lepore, E. (1996). The red herring and the pet fish: Why concepts still can't be prototypes. *Cognition*, 58, 253–270.
- Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 23, 71–87.
- Gagné, C. L. (2000) Relation-based combinations versus property-based combinations: A test of the CARIN theory and the dual-process theory of conceptual combination. *Journal of Memory and Language*, 42(3), 365–389.
- Gagné, C. L., & Spalding, T. L. (2011). Inferential processing and meta-knowledge as the bases for property attribution in combined concepts. *Journal of Memory and Language*, 65, 176–192.
- Gagné, C. L., & Spalding, T. L. (2014a). Conceptual composition: The role of relational competition in the comprehension of modifier-noun phrases and noun-noun compounds. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 59, pp. 97–130). New York: Elsevier.
- Gagné, C. L., & Spalding, T. L. (2014b). Subcategorisation, not uncertainty, drives the modification effect. *Language, Cognition and Neuroscience*, 29(10), 1283–1294.
- Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, 15, 55–71.
- Hampton, J. A. (1991). The combination of prototype concepts. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings*. Hillsdale, NJ: Erlbaum.
- Hampton, J. A., Passanisi, A., & Jönsson, M. L. (2011). The modifier effect and property mutability. *Journal of Memory and Language*, 64, 233–248.
- Jönsson, M. L., & Hampton, J. A. (2006). The inverse conjunction fallacy. *Journal of Memory and Language*, 55, 317–334.
- Jönsson, M. L., & Hampton, J. A. (2008). On prototypes as defaults (comment on Connolly, Fodor, Gleitman, and Gleitman, 2007). *Cognition*, 106, 913–923.
- Jönsson, M. L., & Hampton, J. A. (2012). The modifier effect in within-category induction: Default inheritance in complex noun phrases. *Language and Cognitive Processes*, 27, 90–116.
- Kostelecky, M. (2014). Not induction's problem: Aquinas on induction, simple apprehension, and their metaphysical suppositions. In P. Biondi & L. Groarke (Eds.), *Shifting the paradigm: Alternate perspectives on induction*. Berlin: De Gruyter.
- Laurence, S., & Margolis, E. (1999). Concepts and cognitive science. In E. Margolis & S. Laurence (Eds.), *Concepts: Core readings*. Cambridge, MA: A Bradford Book, MIT Press.
- Levi, J. N. (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.
- Libben, G. (1993) Are morphological structures computed during word recognition?. *Journal of Psycholinguistic Research*, 22(5), 535–544.
- Madden, J. D. (2013). *Mind, matter & nature: A Thomistic proposal for the philosophy of mind*. Washington, D.C.: The Catholic University of America Press.
- Mercier, D.-J. (1950a/1916). *A manual of modern scholastic philosophy (Vol. I)*. St. Louis, MO: B. Herder Book Co.
- Mercier, D.-J. (1950b/1916). *A manual of modern scholastic philosophy (Vol. II)*. St. Louis, MO: B. Herder Book Co.
- Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science*, 12, 529–562.
- Murphy, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language*, 29, 259–288.
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9, 35–58.
- Perler, D. (2001). Essentialism and direct realism: Some late medieval perspectives. *Topoi*, 19, 111–122.

- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus*. New York, NY: Springer.
- Prasada, S., & Dillingham, E. M. (2009). Representation of principled connections: A window onto the formal aspect of common sense conception. *Cognitive Science*, 33, 401–448.
- Pustejovsky, J. (1995). *The generative Lexicon*. Cambridge, Mass.: MIT Press.
- Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, 43, 103–121.
- Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, 59, 413–425.
- Reynolds, P. L. (2001). Properties, causality, and epistemological optimism in Thomas Aquinas. *Recherches de Théologie et Philosophie Médiévales*, 68, 270–309.
- Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, 8, 337–361.
- Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, 12, 485–527.
- Spalding, T. L., & Gagné, C. L. (2013). Concepts in Aristotle and Aquinas: Implications for current theoretical approaches. *Journal of Theoretical and Philosophical Psychology*, 33, 71–89.
- Spalding, T. L., & Gagné, C. L. (2015). Property attribution in combined concepts. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 41(3), 693–707.
- Spalding, T. L., Gagné, C. L., Mullaly, A. C., & Ji, J. (2010). Relation-based interpretation of noun-noun phrases: A new theoretical approach. *Linguistische Berichte, Sonderheft*, 17, 283–315.
- Spalding, T. L., Stedman, J., Hancock, C., & Gagné, C. L. (2014). Intentionality and the Aristotelian-Thomistic view of concepts. *Journal of Mind and Behavior*, 35, 245–262.
- Wippel, J. F. (2000). *The metaphysical thought of Thomas Aquinas: From finite being to uncreated being*. Washington, D.C.: The Catholic University of America Press.
- Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, 4, 167–183.
- Wu, L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, 132, 173–189.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Conceptual Versus Referential Affordance in Concept Composition

Louise McNally and Gemma Boleda

Abstract One of the defining traits of language is its capacity to mediate between concepts in our mind, which encapsulate generalizations, and the things they refer to in a given communicative act, with all their idiosyncratic properties. This article examines precisely this interplay between conceptual and referential aspects of meaning, and proposes that concept composition (or concept combination, a term more commonly used in Psychology) exploits both: *Conceptually afforded composition* is at play when a modifier and its head fit as could be expected given the properties of the two concepts involved, whereas in *referentially afforded composition* the result of the composition depends on specific, independently available properties of the referent. For instance, *red box* tends to be applied to boxes whose surface is red, but, given the appropriate context, it can also be applied to e.g. a brown box that contains red objects. We support our proposal with data from nominal modification, and explore a way to formally distinguish the two kinds of composition and integrate them into a more general framework for semantic analysis. Along the way, we recover the classically Fregean notion of sense as including conceptual information, and show the potential of distributional semantics, a framework that has become very influential in Cognitive Science and Computational Linguistics, to address research questions from a theoretical linguistic perspective.

1 Introduction

The goal of this paper is to confront and explore the larger implications of a problem that we have repeatedly observed in our ongoing work on the semantics of modification within noun phrases, which is one instantiation of concept combination.

L. McNally (✉) · G. Boleda
Department of Translation and Language Sciences, Universitat Pompeu Fabra,
Barcelona, Spain
e-mail: louise.mcnally@upf.edu

G. Boleda
e-mail: gemma.boleda@upf.edu

The problem is that, in the absence of context, sometimes the default interpretation for the modifier-noun combination is so strong as to make other possible interpretations seem impossible, whereas in context any interpretation—even the seemingly impossible—is possible. Here is just one example, involving so-called ethnic adjectives, which provide information about the ethnic origin, nationality or other locational origin of individuals.¹ Kayne (1984) and many others have claimed that when ethnic adjectives like *Canadian* combine with eventuality-denoting nouns, the adjective must contribute information about the most external argument of that eventuality, typically the agent. When it does not, a prepositional phrase expressing the corresponding participant role must be used. Thus, in (1), where the context does not previously mention Yeltsin visiting Canada, the PP *to Canada* rather than the adjective *Canadian* is what the author chose, and indeed the adjective sounds very odd.

- (1) Yeltsin met the prospective Democratic presidential candidate Bill Clinton on June 18. His itinerary also included **an official visit to Canada**??**an official Canadian visit**. (BNC)

However, one does not have to go far to find counterexamples to Kayne's claim. When context or background knowledge make salient that some role other than agent is assigned to the location/ethnicity, the adjective is perfectly felicitous and attested, as in (2).

- (2) Prince Edward and wife begin **Canadian visit**

(<http://metronews.ca/news/canada/365325/prince-edward-and-wife-begin-canadian-visit/>)

Confronted with the contrast between the strong default interpretation and the possibility of any interpretation in context, linguists have tended to follow one of two routes, both of which we will discuss and exemplify below. The first involves taking the default interpretation as the crucial fact to account for, leaving the non-default interpretations in context unexplained. The second involves providing an analysis that is weak enough to capture all interpretations, and saying little or nothing about the strength of the default interpretation. In this paper, we argue that, in effect, both routes must be taken because two fundamentally different interpretative processes can be appealed to in the composition of modified noun phrases or, more generally, in concept composition. Specifically, we take default interpretations to be the result of what we will call *conceptually afforded concept composition*, and non-default interpretations to be the result of *referentially afforded concept composition*. We borrow the term *affordance* loosely from the psychology literature, specifically the interpretation of the term in Chemero (2003), as we discuss in further detail in Sect. 3.1.

¹The sources of examples taken from corpora or the internet via Google searches are indicated in parentheses after the example. 'BNC' refers to a local installation of the British National Corpus (Burnage and Dunlop 1992), though we have also consulted the English language corpora at <http://corpus.byu.edu> for supplementary information.

This distinction builds on the long-standing observation that language mediates between concepts in our mind and the things they refer to in the world (Ogden and Richards 1923, among many others). We take these connections to concepts and to the world to be distinct aspects of language, each of which facilitates a different process of concept composition.² Take for instance the phrase *red box* in the examples in (3). In the absence of any context, *red*, when modifying *box* (or indeed any noun denoting a physical object), refers to its color, and so we can usually paraphrase (3-a) as “Identify a box that is red in color and put the relevant scarf inside it”. However, it may also refer to other properties of the box referent, such as the intended color of its contents, if the discourse context makes the relevant property clear (3-b).

- (3) a. Put the scarf in the red box.
 b. (*Context: For a fundraising sale, Adam and Barbara are sorting donated scarves according to color in different, identical, brown cardboard boxes. Barbara distractedly puts a red scarf in the box containing blue scarves.*)
 Adam: Hey, this one belongs in the red box!

We call cases like (3-a) *conceptually afforded*. In these cases, some component(s) of the concepts contributed by two expressions in a phrase match in a way that indicates how they should be composed, and interlocutors avail themselves of such a suggestion. This matching invites the hearer to identify red as the color of the box in (3-a).³ In contrast, in *referentially afforded* cases like (3-b), specific, independently available information about the referent described by the phrase is used to guide the way in which the concepts in question are composed.

This paper has three goals. First we develop this distinction, which has a precedent in Asher (2011), in an explicit manner and support it with empirical evidence we gathered in previous work. Second, we suggest modeling conceptually-afforded concept composition via (compositional) *distributional semantics*, which represents meaning as a function of the contexts in which words and phrases appear in naturally occurring language data, usually a large text corpus (Landauer and Dumais 1997; Turney and Pantel 2010). We consider this way of modeling concepts to be similar in some of its basic properties to the view of concepts espoused, for example, in Barsalou (2017). A fundamental hypothesis of some work in distributional semantics (e.g. Lenci 2008) is that the resulting semantic representations can be used to model the concepts associated with words. For this reason, we will present a brief introduction to distributional methods in Sect. 4. Finally, we propose a way to formally distinguish the two kinds of concept composition and integrate them into a more general framework for semantic analysis.

²See also Pelletier (2017) for discussion of these two dimensions of meaning in the context of a broader examination of what he calls “Subjectivist” and “Objectivist” approaches, respectively.

³See Hampton (2017) for discussion of experimental work that explores how subjects perform this kind of composition.

2 Two Approaches to Analyzing Modification

We start by discussing previous approaches to the problem outlined in the introduction. As mentioned, the existence of strong, but overridable, defaults in the interpretation of modifiers has led to two lines of analysis. The first involves the proposal of an inventory of primitive semantic relations to capture the defaults; the latter, the use of an underspecified modification relation which gets resolved in context or via an appeal to indexicality. We consider these in turn.

The use of primitive semantic relations to mediate in modification has a long history. We cite two representative and well-known examples here. The first involves Levi's (1978) analysis of relational adjectives such as *microscopic* or *tropical* (ethnic adjectives like *Canadian* are also considered a subclass of the relational adjectives). Relational adjectives (as their name indicates) are morphologically adjectives, but they are also noun-like in several respects: They are synchronically or diachronically derived from nouns; they are generally defined as introducing a relation between an individual of the sort described by the adjective's nominal stem and that described by the modified noun (Bally 1944); and they have a more restricted syntactic distribution than other types of adjectives, occupying in English essentially the same position in nominal syntax as do noun modifiers of nouns, very close to the head noun (e.g. *computer* in *computer store*). Some examples from Levi (1978, pp. 27–28) are provided in (4), with typical paraphrases:

- (4) a. microscopic analysis—analysis carried out using a microscope
 b. tropical butterflies—butterflies found in the tropics
 c. planetary mass—mass of a planet
 d. editorial comment—comment by an editor
 e. dramatic criticism—criticism of drama

Levi proposed that such examples are derived from an underlying structure that makes the relation in question explicit. She further proposed that an inventory of primitive relations could be specified: CAUSE, HAVE, MAKE, USE, BE, IN, FOR, FROM, and ABOUT. For the derivation of examples involving deverbal nominalizations, as in (4-d, e), she proposed somewhat more complex derivations that nonetheless also availed themselves of primitives, including in some cases AGENT and PATIENT.

A second example of an appeal to primitive relations emerged in part from the strong tendencies in the interpretation of (non-relational) adjectives described in Pustejovsky (1995).

- (5) a. red pen—pen that writes in red or that is red on the surface
 b. red apple—apple whose skin is red
 c. quick meal—meal that is quick to eat or to prepare

To account for these interpretations Pustejovsky argues that the lexical entry for content words (including nouns) should include what he called a *Qualia Structure* with four features, each corresponding to a *quale*: FORMAL, CONSTITUTIVE, AGENTIVE, and TELIC. The FORMAL quale characterizes the general ontological properties of an

object; the CONSTITUTIVE, its parts; the AGENTIVE, how it comes into existence; and the TELIC, its function. Pustejovsky proposes that adjectives can restrict the denotation of a noun by placing conditions on the values of the different qualia in the noun's semantic representation.

The logical representations in (6) illustrate how this approach can be used in the sorts of modification that interest us here. In (6-a), the primitive AGENT specifies the relation between Canada and the visit; in (6-b), the primitive CONSTITUTIVE acts as an operator on the representation of *apple* to retrieve indirectly a part of the apple to which the property denoted by *red* can be ascribed.

- (6) a. *Canadian visit*: $\lambda e[\mathbf{visit}(e) \wedge \mathbf{AGENT}(e, \mathbf{Canada})]$
 b. *red apple*: $\lambda x \exists y[\mathbf{apple}(x) \wedge \mathbf{CONSTITUTIVE}(\mathbf{apple}) = \mathbf{PART-OF}(y, x) \wedge \mathbf{red}(y)]$

The use of semantic primitives to capture modification relations has two main advantages. First, it speaks to the very strong intuitions that the literature reports about default or productive interpretation processes (see e.g. Levi 1978, pp. 84–86). Second, similar defaults are observed cross-linguistically—for example, Pustejovsky's theory has been applied to various languages, and Levi observes that she found evidence for a similar set of primitives in a study of Modern Hebrew (Levi 1978, p. 86). Clearly, there is something to be captured in these data.

However, the use of primitives of any sort, at least as the only compositional strategy, has also long been argued to be problematic. On the one hand, it is clearly too strong insofar as no set of necessary and sufficient primitives can be provided to account for all cases.⁴ Levi herself observes (p. 84; see also p. 238ff.) that the goal of her study is to account for patterns of modification that are *productive*, as opposed to *possible*: In other words, her aim was a theory of why, even if we *can* interpret, for example, a phrase such as *Korean passengers* as 'passengers on Korean Airlines',⁵ our first inclination is arguably not to do so but rather to interpret it as 'passengers from Korea'. On the other hand, the use of primitives is too weak. As e.g. Clark (1992) and Murphy (2002) observe, even when such primitives might apply, they are insufficiently granular: There are cases in which they provide too little information about the exact nature of the relation instantiated by any given primitive. This is already apparent in the analysis of *red apple* in (6). The CONSTITUTIVE quale introduces a part of the apple, but it does not specify which part, and so the inference that it is the skin of the apple (or, more generally, its surface) is not directly accounted for. A representative example involving a relational adjective is *an electrical fire*, which could be paraphrased as 'a fire caused by electricity': This case is even more problematic than the *apple* example insofar as the paraphrase does not capture the fact that the term is used to refer to fires caused by malfunctions in electrical systems and not, for example, by lightning.

⁴Observations to this effect with respect to productive compounding, which shares important properties with modification of the sort discussed by Levi, can be found as far back as Jespersen (1942). See Gagné et al. (2017), Westerlund and Pylkkänen (2017), and references cited in these works for further discussion of the complexities involved in modification.

⁵Example taken from http://www.koreatimes.co.kr/www/news/biz/2014/09/373_135962.html.

Given these problems, a second line of approach to modification has involved sacrificing the coarse generalizations embodied in primitives in favor of broader empirical coverage. On one version of the approach, modification is mediated by a maximally underspecified relation whose value, much like that of a pronoun, is resolved in context (examples include McNally and Boleda 2004; Kennedy and McNally 2010). On another (see e.g. Bosch 1983; Rothschild and Segal 2009) adjectives have as their lexical content functions from contexts to contents, that is, Kaplanian characters (Kaplan 1989). These analyses are respectively illustrated in (7).

- (7) a. *Canadian visit*: $\lambda e[\mathbf{visit}(e) \wedge R_i(e, \mathbf{Canada})]$
 b. *red apple*: $\lambda x[(\mathbf{red}_i(\mathbf{apple}))(x)]$

Again, this approach has both advantages and disadvantages. On the positive side, it is appropriately flexible: There is no interpretation that cannot be accommodated under such an analysis. However, its flexibility is arguably also a disadvantage: It has nothing to say about the strength of default interpretations or the fact that we tend to generalize them to new examples (a point made, as noted above, by Levi). Moreover, these analyses have provided no substantive theory of how context intervenes to yield the interpretations that arise.

We know of only one explicit proposal that contemplates the possibility of combining these two general approaches to resolving modification, namely that in Asher (2011). Asher combines a classical, model-theoretically interpreted intensional logic with a separate, proof-theoretic logic of types that is intended to mirror language users' systems of concepts. The latter is used to compute and resolve the basic relations between predicates in composition—for example, it will allow us to determine that, in principle, it must be possible to infer that *red* picks out a type that, when combined with the type picked out by *pen*, yields a type that corresponds to a pen that writes in red.⁶ Though he does not develop the possibility in detail, he suggests (p. 226) that Pustejovsky's qualia could be introduced into his system as type-shifting operators that mediate in this process: For example, WRITE could be the output of a general type coercion operator TELIC applied to the type *pen*, and this information could then be exploited in the semantic composition process. In addition, alongside the possibility of such operators, Asher's system contemplates the possibility of contextually-valued type coercion operators for cases where the discourse structure makes it clear that default value operators such as TELIC would not apply.

The proposal we develop in the rest of this paper shares with Asher's the intuition that there are (at least) two distinct sorts of composition processes involved in computing the interpretation of a sentence. Our contributions will consist in laying out the proposal in more explicit terms, providing new empirical support for this dual system, the use of distributional semantics an alternative to Asher's logic of types, and a specific proposal for formalizing the distinction using Discourse Representation Theory (Kamp 1981).

⁶Asher's logic of types builds on an approach to type theory, now sometimes referred to as "Modern Type Theory", developed within intuitionistic logic by Martin-Löf (1984). Due to space limitations, we refer the reader to Asher's book and Chatzikyriakidis and Luo (2017) for details.

3 A Dual System for Semantic Composition

3.1 *Conceptually Versus Referentially Afforded Composition*

We begin with a very programmatic proposal concerning two ways in which the construction of meaning can be mediated. Our proposal is based on the following assumption.

Assumption The construction of meaning draws on connections we make between linguistic expressions and our conceptual structure, on the one hand, and the world, on the other.

This assumption is of course familiar from traditional semiotic models and also resonates with the “dual content” model recently proposed in Del Pinal (2015), which provides (p. 44ff.) a useful overview of the different ways in which language, conceptual structure, and the world have been related to each other both in the philosophy of language and cognitive psychology literature. The assumption also underlies the classic Fregean model that distinguishes *sense* (*Sinn*), which Frege suggests forms part of the ‘common treasure of thoughts that [humanity] transmits from one generation to another’, and *reference* (*Bedeutung*) (Frege 1892, p. 29⁷). However, in modern formal semantics in the Montagovian tradition, despite its Fregean roots, conceptual structure has largely been set aside. In this latter tradition, Fregean sense has largely been substituted for the notion of *intension*, modeled non-psychologically as, for example, a function from possible worlds to truth values.

We recover the classically Fregean notion of sense as including conceptual-like information and propose that both conceptual and referential aspects of meaning play a role in composition. Specifically, we can think of them as affording concept combination in different ways. Our use of the term *affordance* is based on Chemero’s (2003) development of the notion, originally due to Gibson (1979); it is also inspired in Rietveld’s (2008) extension of the notion to higher cognition. Chemero defines affordance as a relation between features of situations and abilities of organisms, and argues that to perceive an affordance is to recognize that the feature in question facilitates an action by the organism. The classic example is a mug with a handle: If a person who has never seen a mug gets to interact with it, it is very likely that she will grab it by the handle. The mug, by its shape, affords the grabbing-by-the-handle action on the part of the person.

Our extension of this idea to the case of language is very simple. We take the connection to concepts, on the one hand, and to the world, on the other, to be distinct features of language, each of which facilitates—that is, affords—a distinct composition process. If we posit that language users have access to both of these features and the corresponding processes that they facilitate, the tension we observed between default and highly context dependent interpretations in Sect. 2 disappears.

⁷We cite the translation by Max Kölbel, published in Byrne and Kölbel (2009) which includes references to the original pagination by Frege.

The default interpretations, we argue, can be understood as the result of conceptually-afforded concept composition. These are the interpretations that are immediately available in the absence of discourse context, and they are productive, suggesting that they build on regularities in our lexical knowledge—that is, the connections between words and concepts. For instance, the fact that physical objects typically have colored surfaces will afford the interpretation of a color term modifying a noun denoting a physical object as describing surface color, as in *red box* (see example (3-a)) and *red apple* (5-b). The fact that a writing instrument produces text or images with a particular color, and that this color may vary from one writing instrument to the next—part of our concept of what a writing instrument is like—affords the interpretation of *red pen* as a pen that writes in red (example 5-a). Note that the use of a color term with *pen* is easily extended to other writing instruments with the same general properties, such as *pencil*, *crayon*, or *marker*. Similarly, the fact that analyses are carried out using instruments, and that microscopes are instruments, affords the interpretation of *microscopic analysis* provided in (4-a) above. Different species of animals tend to require different climates, so again, the interpretation of (4-b), with *tropical* describing a climate, is on our view conceptually afforded.

Note that these interpretations arise from very detailed conceptual knowledge, presumably accessible from the words involved. The primitive-based analyses discussed in the previous section are too coarse-grained to allow for these interactions; the lack of conceptual information in typical formal semantic approaches doesn't allow for them either. Thus we need a richer and more nuanced lexical representation; in Sect. 4 we show how distributional semantics can serve this purpose.

The notion of conceptual affordance also allows us to make predictions about combinations of modifiers and nouns that will be infelicitous out of the blue. Interestingly, Vecchi et al. (2011) developed a computational model (using, it is worth noting, distributional semantics) that was able to partially distinguish between (out of the blue) deviant versus possible adjective-noun phrases. Vecchi et al. randomly selected a set of phrases that were unattested in a very large corpus and tested whether their model would group them in ways that correlated with whether or not the phrase was acceptable to human judges. Examples of unattested but semantically acceptable phrases included *vulnerable gunman*, *huge joystick*, and *blind cook*; deviant phrases included, for instance, *blind pronunciation*, *parliamentary potato*, and *sharp glue*. The acceptable phrases are similar to those we have hypothesized above to involve conceptually afforded composition. For instance, joysticks are physical objects and have a size dimension that can be modified by *huge*. In contrast, it is not obvious in the absence of a specific context along which conceptual dimension a pronunciation could be blind, what kind of relation might exist between potatoes and parliaments, or what it would mean for glue, which is not rigid, to be sharp.

Now, it is possible to find a semantic interpretation for the allegedly deviant phrases. For instance, imagine that potatoes were thrown at parliamentary members in a protest concerning the recent economic crisis in Spain, and that one of the potatoes knocked out the president of the parliament and was retrieved and put on display. This object could well be dubbed *the parliamentary potato*. We submit that

such interpretations are the result of referentially-afforded concept composition: they are retrievable only once we have a specific candidate (or small set of candidates) for who or what is being referred to with the phrase, along with a salient set of candidate properties that could be described by the modifier.

To further illustrate referentially-afforded composition, let us return to the use of “red box” in (3-b) and the “Canadian visit” example in (2). In the first case, the situation presents the hearer with two brown cardboard boxes. The speaker can assume that the hearer knows that the boxes each have a context-specific property of being destined to hold objects of a specific color. The use of *red* to modify the *box* color in this case is incongruent with what we can arguably consider the basic concept associated with *box*—the concept cannot afford any meaningful interpretation of the modifier—but the box referents and their context-specific properties can. In the case of (2), recall that the problem is that ethnic adjectives tend to express the agent when combined with eventive nouns. Thus, by default we expect *Canadian visit* to describe a visit made by Canadians. However, in (2) the interpretation on which Canada is the location visited is afforded by specific information about individuals in the context, namely, that Prince Edward and his wife are members of the British royalty, that *Canada* denotes a place, and that Prince Edward and his wife are the agents of the action of beginning a (Canadian) visit.

In order for the distinction between conceptually and referentially afforded concept composition to have bite, we should have independent criteria for identifying the components of the specific concepts being combined. For now, we limit ourselves to the claim that once such criteria are established, it should be possible to predict when a combination of modifier and noun is easily interpreted in the absence of a specific discourse context.

3.2 *Empirical Evidence Supporting the Distinction*

Despite the caveat made in the preceding paragraph, we have been encouraged by the fact that the distinction between conceptually and referentially afforded concept composition gives us insight into puzzling data that we gathered in previous work and for which we had no explanation at the time. We now summarize these data and explain how our proposal predicts them.

First, the modification data reviewed so far point to the fact that modifier-noun combinations can have very plastic interpretations. Our proposal suggests that a large part of this plasticity corresponds to referentially afforded composition. This prediction is supported by empirical data we gathered about relational adjectives, which we introduced in Sect. 2.

As noted above, relational adjectives are typically denominal and, crucially, the adjective-forming morphology has been claimed to be essentially transparent (e.g. Spencer 1999). The only contribution of the adjectival morphology, then, would be

to make explicit that there is some relation between the referent of the noun from which the adjective is derived and the referent of the noun that the relational adjective modifies.⁸

By hypothesis, relational adjectives provide a way to pick up on, in a maximally condensed fashion, the myriad possible relations between the referent of the modified noun and the referent of the adjective's nominal stem (e.g. 'Canada', in the case of *Canadian*). These relations can be identified thanks to general knowledge (e.g. *national anthem*) or be inferred from the meaning of the modified noun (particularly when the noun is deverbal, as in *chemical reaction*); however, we posit that in many cases the relations are in fact afforded by specific information we have about the referents in question in the discourse (e.g. *Canadian visit* in example (2) above).

Boleda (2007) reported that, in Catalan, relational adjectives appear much more often in definite noun phrases than do other types of adjectives: Specifically, in an analysis of a 16.5 million-word, balanced Catalan corpus, relational adjectives appeared almost 60% of the time in definite noun phrases (59%, with a standard deviation of 15%), while other types of adjectives did so a little over 30% of the time. Definite noun phrases are used to refer to individuals that are familiar either from the context or from prior discourse, and referentially afforded concept composition is only possible when the referent is known. Thus, the high proportion of uses of relational adjectives inside definite noun phrases suggests a tendency towards referential affordance in the composition of relational adjectives and nouns.⁹ Without a distinction between the two kinds of composition, it is far from clear how to account for the data in Boleda (2007).

Boleda and colleagues provided more data in the same direction in a statistical study of the British National Corpus (Boleda et al. 2012). The study compared nominal modification using ethnic adjectives (*Canadian*) to modification using prepositional phrases (*from Canada, to Canada, etc.*).¹⁰ The two types of expressions often seem synonymous: For instance, both *Canadian visit* and *visit to Canada* could be used in example (2). However, the results showed that ethnic adjectives are used especially when the discourse makes the semantic relationship between the head noun and the adjective explicit, that is, in contexts where previous information about the referent is available. Factors correlated with the use of these adjectives in the corpus (as opposed to their prepositional phrase counterparts) included, again, the definiteness of the DP containing the ethnic adjective, and also others like the occur-

⁸This claim does not exclude the possibility that, over time, a relational adjective might come to be associated with a specific property or properties, such as the systematic use of ethnic adjectives to pick out characteristic properties of the individuals of a particular origin, as in (i).

(i) Park's parents immigrated to the United States in the 1950s. (...) Park says, "My parents thought the best way to help us succeed was to become very American...".
(W. D. Thomas, *Korean Americans*, Benchmark Books, 2009, p. 59).

⁹Of course, that does not imply that other types of adjectives cannot enter into referentially afforded composition (cf. the *red box* example in (3-b)), but relational adjectives, because of their properties, are expected to do so more often.

¹⁰Recall that ethnic adjectives are usually considered a subclass of relational adjectives.

rence of *visited Canada* before *Canadian visit* in a given discourse. To summarize, both studies constitute evidence for referentially afforded concept composition and show some effects that the use of this composition strategy by language users has on their linguistic output.

A second piece of evidence concerns another prediction of our proposal, namely, that the more context dependent (or referentially afforded) concept composition is, the more difficult it can be expected to be to reconstruct out of context. The results of a study involving computational modeling of adjectival modification, reported in Boleda et al. (2013), are in line with this prediction.

Boleda et al. (2013) used computational semantic methods to produce meaning representations for adjective-noun phrases. They built representations for phrases like *former commentator* in two ways. On the one hand, they constructed a representation of the entire phrase directly from linguistic data, extracting statistics from a large textual corpus. We will call this representation the *observed* representation. On the other, they combined the representations for the individual words in the phrase (also obtained from a corpus) using a computational algorithm. For example, this algorithm took the representation for *former* and that for *commentator* to build a semantic representation for *former commentator*. Boleda et al. then compared this “artificial” or *predicted* representation with the observed one, to see how accurate the prediction was.¹¹

Their results showed that the more typical the property denoted by the adjective is for the entity described by the noun, the easier it is to model the semantics of the phrase. Specifically, the resulting predicted representation of the phrase is more similar to the observed one when the property is more typical. For instance, *former* can be said to be a typical attribute for role-denoting nouns such as *commentator*, *father-in-law*, *teacher*, or *president*, insofar as the concepts associated with these nouns arguably include the specification that the role has a potentially limited duration. Information about duration is supplied by adjectives like *former*, *current*, or *future*. And indeed, the predicted representation for *former commentator* obtained by combining the representations for its two component words was very similar to the observed one. In contrast, the representation for phrases like *former colour* was more difficult to predict from the component words alone: *Colour* does not denote a concept with a clear temporal specification, and the relationship between *former* and *colour* will depend on the object whose colour is being referred to.¹²

We conclude that modification of nouns by adjectives describing typical attributes corresponds to conceptually afforded composition, and at least some uses of atypical modifiers correspond to referentially afforded composition. In the latter cases, without additional evidence from the specific discourse context it is hard to make

¹¹More details on the methods used in this study will be provided in Sect. 4.

¹²Typicality correlates positively with frequency, but the contrasting effects we mention are attested with phrases of similar frequency, including the ones used in the text. Thus, the results of this study cannot be explained simply as a byproduct of the frequency of the phrases in question.

sense of the semantic relationship between the adjective and the noun.¹³ Therefore, the distinction between conceptual and referential affordance in modification helps explain the results in Boleda et al. (2013).

Classical formal analyses of semantic composition involving adjectives are not well-equipped to take into account the degree of fit or typicality relation between the property denoted by the adjective and general features of the concept associated with the noun. Theories like Pustejovsky's Generative Lexicon were designed to do this to some extent; however, as noted in Sect. 2, such theories cannot help with highly context-dependent meaning relations. Thus, the challenge is to find a way to incorporate the distinction between conceptually and referentially afforded concept composition into semantic theory, so as to broaden the theory's empirical coverage. As a first step in addressing this challenge, we turn to distributional semantics.

4 Conceptually Afforded Composition with Distributional Semantics

We propose distributional semantics as a framework to account for conceptually afforded composition because we do not consider other approaches (e.g. standard formal semantics or primitive-based approaches such as the Generative Lexicon) to offer a rich enough representation of a word's meaning to account for the range of effects discussed. However, as it is beyond the scope of this paper to exhaustively compare these different approaches, we limit ourselves here to simply providing enough background on distributional semantics for the reader to be able to follow the formalization presented in the next section, leaving more thorough discussion for future work.

Distributional semantic analyses (Landauer and Dumais 1997; Turney and Pantel 2010; Erk 2012) represent the semantics of a word as a function of the contexts it occurs in. Context can be defined in various ways, but the most typical approach is to define context as the words surrounding the target word in a corpus. A distributional representation for a word will then be a list of context counts, aggregated over the whole corpus and suitably transformed, that is, a *vector*. Figures 1 and 2 offer a toy example. Figure 1 is intended to illustrate how even a small context window reveals repeated examples of co-occurrences between a target word (here, *moon*) and other words that are suggestive of our knowledge about the target. Figure 2 exemplifies a partial vectorial representation for the words *moon*, *sun*, and *dog*.¹⁴ The vectors show how the distributional representation mirrors some semantic similarities and differences between these words: All three can appear with *shadow*, but, while *moon* and *sun* appear with words such as *planet* or *shine*, *dog* does not. *Moon* and *sun*

¹³Note that the semantic representations used are aggregates over all uses of a word in a corpus and do not reflect discourse structure.

¹⁴Figures 1, 2, 3 and 4 are adapted from materials by Marco Baroni.

Fig. 1 The basic data for distributional semantic representations: contexts

curtains open and the	moon	<i>shining</i> in on the barely
rough the <i>night</i> with the	moon	<i>shining</i> so <i>brightly</i> , it
surely under a <i>crescent</i>	moon	, thrilled by ice-white
is dazzling snow, the	moon	has <i>risen</i> <i>full</i> and <i>cold</i>
and the <i>temple</i> of the	moon	, driving out of the hug
in the <i>dark</i> and now the	moon	<i>rises</i> , <i>full</i> and amber
bird on the <i>shape</i> of the	moon	over the <i>trees</i> in front

Fig. 2 Semantic representation: vectors of context counts

	<i>planet</i>	<i>night</i>	<i>full</i>	<i>shadow</i>	<i>shine</i>	<i>crescent</i>
moon	10	22	43	16	29	12
sun	14	10	4	15	45	0
dog	0	4	2	10	0	0

are similar in representation, but not identical: for example, *full* and *crescent* occur primarily with *moon*, while *shine* is a more typical context for *sun* than for *moon*.

A vector for a word as used in distributional semantics ranges from a few hundred to a few thousand dimensions (that is, contexts or transformations thereof), thus providing a very rich, flexible representation for word meaning. However, this makes it difficult to inspect it manually.¹⁵ The power of distributional semantics lies in its use of well defined linear algebra techniques to manipulate these vectors, yielding useful information about the semantics of the involved words. We visualize one kind of technique used in Fig. 3, where simple, two-dimensional vectors for the words *moon*, *sun*, and *dog* are visually represented. The two dimensions depicted in the graph (corresponding to word contexts) are *shadow* and *shine*, with the values shown in the left part of Fig. 3. The geometric distance (e.g., the Euclidean distance; see discontinuous lines) between the vectors for *moon* and *sun* is smaller than the distance between the vectors for *moon* and *dog*. Crucially, the algebraic techniques that we can visualize with two dimensions generalize to any number of dimensions. Thus, in distributional semantics, geometric distance corresponds to semantic distance.

Distributional semantic methods are highly successful at modeling word meaning because they are based on linguistic data naturally produced by humans, as manifest in large text corpora drawn from the internet and other sources. The representations are rich, with hundreds or thousands of dimensions providing different bits of contextual information. Also, distributional representations are naturally graded; for instance, the notion of semantic distance is a continuum, with words being more or less distant. This makes them useful for semantic phenomena such as the typicality effect observed in the previous section.

Recently, researchers have begun to explore *compositional* distributional semantics, giving a distributional representation not only to words but also to phrases and

¹⁵The context counts in a real distributional model are also typically further operated upon to remove noise in the data and make them more compact; see the references in this section for more information. Also, recently, neural network or deep learning models have been shown to outperform traditional count-based methods such as the ones explained in this section (Baroni et al. 2014b). Since nothing we say in this paper hinges on the particular type of model chosen, for clarity we present count-based models only.

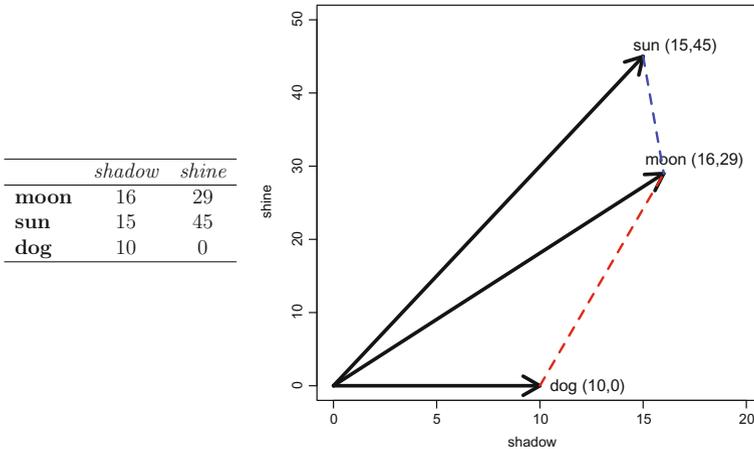


Fig. 3 Semantic distance as geometric distance

Fig. 4 Vector composition by addition

	<i>planet</i>	<i>night</i>	<i>blood</i>
red	15	3	21
moon	24	15	1
red + moon	39	18	22

even sentences (Mitchell and Lapata 2010; Coecke et al. 2011; Socher et al. 2012; Baroni et al. 2014a; Pham 2016 among many others); the previous work we presented at the end of Sect. 3.2 falls into this line of research. Here, the challenge is typically framed as capturing how composition changes the values of the vectors. For instance, *blood* is not a relevant context for *moon*, but when *red* modifies *moon* it does become relevant (see Fig. 4). This kind of effect is achieved by applying composition operations to build the meaning representation of the phrase from the representations of its constituents. A very simple but stubbornly effective method is to simply add up the word vectors, as in Fig. 4 (Mitchell and Lapata 2010), but more sophisticated methods have been designed that sometimes yield better results (Baroni and Zamparelli 2010). Nothing we say in this paper depends on the chosen method for composition, hence we will simply use $\text{comp}(\overrightarrow{\text{red}}, \overrightarrow{\text{moon}})$ for the distributional representation of the phrase *red moon* obtained by applying a composition function to its constituent word vectors, $\overrightarrow{\text{red}}$ and $\overrightarrow{\text{moon}}$ (we represent word vectors with an overhead arrow).

Note, finally, that there is an alternative method for obtaining a distributional representation for a phrase, namely, to directly extract it from the corpus, just as representations for words are generated (Fig. 5). Because it is based on counts for actual occurrences of phrases in corpora, this representation should be a faithful rendering of the meaning of the phrase, and this is why we used it as a benchmark in the research in Boleda et al. (2013), to compare to the result of compositionally obtained

<i>a large</i>	red moon	,	Campana			
<i>a blood</i>	red moon	hung over				
<i>glorious</i>	red moon	turning				
The <i>round</i>	red moon	,	she 's			
<i>a blood</i>	red moon	emerged			<i>planet</i>	<i>night</i>
rains ,	red moon	<i>blows,</i>			34	20
<i>monstrous</i>	red moon	had climb			red moon	31
A very	red moon	<i>rising</i> is				
under the	red moon	<i>a vampire</i>				

Fig. 5 Corpus-extracted distributional representation for phrase *red moon*

(predicted) vectors. However, this technique can only be used for sufficiently frequent phrases. Since of course many possible phrases will not occur frequently or even at all, composition is still needed to build a representation for many phrases.

Because of their data-driven nature and their rich representation of meaning, compositional distributional representations for phrases are able to account for subtle nuances of meaning arising from the combination of modifiers and nouns. For instance, Baroni and Zamparelli (2010) report that the most similar element in a large semantic space to the phrase *historical introduction* is *historical background*; to *small drop, droplet*; to *common understanding, common vision*. Though crude and incomplete as an approximation of what concepts are (as the discussion in Barsalou, 2017, will make apparent), these representations have the advantage of being easy to construct and incorporable into a testable interpretive model. We therefore adopt them for modeling conceptually afforded composition in the formalization we offer in the next section.

5 A Mixed Model for Two Types of Semantic Composition

We next sketch how the mechanisms of conceptual and referential affordance can both be incorporated into a single, mixed interpretive model (see Boleda and Herbelot 2016 for a review of previous work combining formal and distributional semantics).¹⁶

We will use Discourse Representation Theory (DRT) as the scaffolding for our semantics. We use DRT because (1) the notion of discourse referent is crucial for implementing referentially afforded composition, and (2) the most recent research on compositional distributional semantics has not yet been able to show how such

¹⁶Some of the basic discussion in this section concerning the integration of distributional semantics and formal semantics is drawn from McNally (2017); however, the application to DRT is new here, as is the idea of referentially-afforded composition. See Garrette et al. (2011) for a different approach to combining distributional semantics and Discourse Representation Theory, and see Pelletier (2017), for extensive discussion of “two-tiered” semantic theories, of which he takes this proposal to be an example.

models can provide effective analyses of referential grounding or discourse dynamics (Bernardi et al. 2015; Sadrzadeh and Purver 2015). This latter state of affairs leads us to tentatively hypothesize that compositional distributional semantics is best used to model only those parts of semantic composition that are, in our terms, conceptually afforded.

For reasons of space, we must assume basic familiarity with DRT; the reader is referred to e.g. Kamp (1981) or Kamp and Reyle (1993) for background. Our implementation of DRT will be entirely standard, with just three exceptions. First, we need a means of connecting distributional semantic representations to Discourse Representation Structures (DRSs). Second, as a result of doing this we will introduce minor modifications in our treatment of nominal and adjectival predication with respect to what is more generally assumed. Finally, we will need a way to distinguish conceptually afforded from referentially afforded composition.

We incorporate distributional semantics by building on the idea in Zamparelli (1995) that nouns (and not just certain kinds of generic noun phrases) denote Carlsonian kinds (Carlson 1977).¹⁷ The crucial step is to use distributional semantic representations rather than atomic abstract entities as models for kinds. However, as with the classic treatment of kinds as abstract entities, these distributional representations will be coded in the DRSs as constants. Since distributional representations are, mathematically-speaking, vectors, the constants we use for them will be indicated with an overhead arrow (e.g. $\overrightarrow{\text{box}}$), as noted in the previous section.

We further extend Zamparelli's idea to adjectives, also interpreting them as vectors (e.g. $\overrightarrow{\text{red}}$). Since adjectives are not assumed to denote natural kinds but rather to pick out properties, this proposal can be seen as generalizing Zamparelli's, substituting concepts for both kinds and properties along the way. Distributional vectors will thus serve as very crude representations for concepts.¹⁸ The crucial step will be to allow the distributional representations for nouns and adjectives to combine with each other to yield new representations of the same type, whose role in the DRT part of the semantics is exactly analogous to the role of the representations for unmodified nouns.

In the previous section we briefly sketched how the composition of two vectors works. We assume that the grammar of a language indicates when semantic composition for certain phrases involves the composition of vectors, as opposed to other

¹⁷Zamparelli posited that nouns come to denote sets of entities only in the semantic composition process, and used type-shifting operations licensed by functional morphosyntax to do this. See below.

¹⁸We should insist that we are not making anything like the claim that concepts, whatever they are, consist only of distributional information, even if such information may play a role in concept formation. Rather, we are using distributional representations to model concepts primarily because they have certain properties that we hypothesize concepts to share and because they have various attractive methodological features, such as that of allowing us to make testable predictions of various kinds.

Note also that, to our knowledge, the formal semantics literature has largely avoided the deeper question of what kinds are and how they relate to concepts. Our reinterpretation of Zamparelli should not be viewed as reflecting any particular position on how kinds have been understood in this literature. See McNally (2017) for further discussion.

sorts of semantic operations. The composition of vectors happens outside of the DRT model, but as the result is also a vector, it can, like the component vectors, be associated with a constant in a DRS, which we will represent as e.g. $\overline{\text{comp}(\overline{\text{red}}, \overline{\text{box}})}$. In other words, constants of the form $\overline{\text{red}}$, $\overline{\text{box}}$, and $\overline{\text{comp}(\overline{\text{red}}, \overline{\text{box}})}$ are all of the same type. Thus, distributional semantics will give us a relatively concrete algebraic model for simple and complex concepts on which both sorts of concepts are of fundamentally the same nature, much in the way lattice-theoretic structures serve as models for treating atomic entities and pluralities as fundamentally similar types of objects (Link 1983).

The next piece we need is a way to exploit nouns and adjectives with such interpretations within DRT, so that referents can be associated with the concepts that nouns and adjectives pick out. Zamparelli used Carlson's (1977) realization relation, which we represent here as **Realize**, aims to do this: This relation holds between an object and a kind just in case the object constitutes an instance of the kind.¹⁹ Again following Zamparelli, we assume that the **Realize** relation is introduced by (possibly abstract) functional morphosyntax that turns a noun into an expression that denotes a set of entities. As a first approximation, then, we can represent a referential expression such as *a box* as in (8), where u is the discourse referent introduced by the phrase, which must satisfy the condition that it is a realization of the concept $\overline{\text{box}}$.

(8)

u
Realize ($u, \overline{\text{box}}$)

Now consider modification. Prior to the point in the syntax at which the **Realize** relation is introduced, the composition operations at work will combine vector-denoting expressions; this corresponds to concept composition. We model conceptually-afforded composition as the result of composing adjective and noun vectors directly into a new vector, corresponding to a complex concept, which can then stand in the **Realize** relation to a discourse referent, as in (9).

(9)

u
Realize ($u, \overline{\text{comp}(\overline{\text{red}}, \overline{\text{box}})}$)

The syntactic rules of the language will have to make it clear when this sort of composition can be appealed to and when not; interestingly, studies of the syntax of modification clearly indicate that syntax could, indeed, encode this kind of information (see, e.g. McNally and Boleda 2004 and Bouchard 2005 on adjective ordering constraints of the sort exemplified by relational adjectives).

Now let us consider referentially-afforded concept composition. As mentioned in Sect. 3, this is attested only when the referent of the nominal is already familiar in the

¹⁹Carlson's ontology also included stages (spatiotemporal slices) of individuals, which could also instantiate kinds, but these will not play a role in our discussion.

discourse. This referent therefore plays a role in the interpretation of the combination of the modifier and noun. We see two ways in which this could be implemented. One would be to take the referent to modulate the composition operation that combines the adjective and noun vectors. This could be represented as in (10), where the subscript u indicates modulation by referent u .

(10)

$$\frac{u}{\mathbf{Realize}(u, \mathbf{comp}_u(\overrightarrow{\mathbf{red}}, \overrightarrow{\mathbf{box}}))}$$

On this view, the concept associated with *red* in the context would be exactly the same across all contexts, but its interaction with the concept contributed by the noun would vary from one context to the next, for example by the use of varying weights on the sums or products of the vectors.

Alternatively, the vector corresponding to the adjective could be modified as a function of the referent, i.e. reinterpreted as an ad hoc, referent-mediated property, as could be represented in (11).

(11)

$$\frac{u}{\mathbf{Realize}(u, \mathbf{comp}(f(u, \overrightarrow{\mathbf{red}}), \overrightarrow{\mathbf{box}}))}$$

On this view, the composition operation as such is not altered in any way; rather the input to that operation is. In other words, *red* in this example would simply be associated with a different concept in the context in question. Further research will be needed to determine which of these options constitutes a better analysis of the facts, or, indeed, if they are empirically distinguishable. However, it is worth noting that this latter approach closely resembles the indexical interpretations of adjectives proposed by Bosch (1983) and Rothschild and Segal (2009), briefly introduced in Sect. 2.

These analyses do not offer an account of *how* context intervenes to determine the referentially afforded interpretation; in this, unfortunately, we are in good company, as no theory we know of offers such an account, and the area is one in which much more research is needed.

We close this section with some very brief, speculative comments on how the proposed analysis relates to classical analyses of adjective modification of nouns within formal semantics. Such modification has been analyzed in two ways: Either by treating the adjective as a second-order property that takes the noun as its argument, or by treating it as a first-order property that is combined via conjunction or set intersection with the (first-order) property denoted by the noun (see e.g. Kamp 1975; Siegel 1976; Larson 1998, among many others, for proposals and discussion). The latter analysis is appropriate specifically for cases of so-called intersective modification, when the adjectival and nominal properties are each entailed to hold of the individual being described. The former is more general and can be used not only

for intersective modification but also for non-intersective modification, namely, subjective modification, where the adjectival property does not obviously directly hold of the referent but the noun property does (cf. *molecular biologist*), and intensional modification, where the nominal property is not entailed to hold (or is entailed not to hold) at the time or world of ascription (*former mayor* or *alleged thief*).

All of the above-sketched implementations of concept composition are counterparts of non-intersective modification. In no case is the concept contributed by the adjective directly related to the referent. Moreover, as we have set things up, our analysis of concept composition directly captures the intuition developed in Landman (2001) and Partee (2010) that all adjective-noun combinations, even intensional modification, are, in some sense, subjective, that is, the nominal description is always somehow used to identify the referent, insofar it contributes positively to the eventual complex description that the referent is related to via the **Realize** relation. Of course, it remains to explore how to reproduce the entailment effects of the world and temporal parameters that have played a role in traditional analyses of intensional adjectives, but we note that one surprising result of the study in Boleda et al. (2013) was that intensional adjectives turned out to be no more difficult to model in distributional semantics than other kinds of adjectives, insofar as, all other things being equal, compositional distributional semantic techniques could predict the semantic representation for phrases containing intensional adjectives from the representations of the component words just as well as they could for phrases containing nonintensional adjectives (see Sect. 4, above).

6 Conclusions

Semantic composition is a dynamic process that cannot be understood without simultaneously considering what we are referring to and the concepts associated with the words we are using. Concepts, and thus the words associated with them, encode significant regularities. At the same time, they are plastic, insofar as we must use a finite vocabulary to describe a potentially infinite variety of situations and generalizations in the world. However, once a word is applied to a referent, that word is grounded in a very specific manner, and the referent can influence the way we understand the word and its associated concept(s) in the context of use. This interplay between our conceptual structure and the world is what motivated the first contribution of this paper, namely to propose that modification works in two ways: It can be conceptually afforded, when the modifier and the head introduce concepts that fit to form a complex concept, and the speaker and the hearer use this fit in their interpretations; or referentially afforded, when the result of combining the modifier with the noun depends on specific properties of the referent. This proposal has an antecedent in Asher (2011), but we have made it more explicit and have proposed a specific analysis combining distributional semantics and DRT.

Along the way we hope to have made a case for further exploring distributional representations within semantic theory. They are automatically induced (and thus

easy to construct and empirically well-founded), have some psychological plausibility (Landauer and Dumais 1997 and subsequent work), and offer a wealth of empirical data. Distributional semantic representations also avoid some of the weaknesses of semantic primitives: Since they generally encode a relatively large number of features with continuous values,²⁰ they can express many more nuances of meaning than a small set of discrete features, while at the same time accounting for default interpretations. The key is to recognize their limitations. In this respect, we consider promising the division of labor between distributional semantics and a referential semantic framework like DRT.

Acknowledgements We are grateful to Marco Baroni, who was involved in an earlier phase of this research, for his input. Also thanks to the audience at the Utrecht Concept Composition and Experimental Semantics/Pragmatics Workshop, the editors of this volume, Galit Weidman-Sassoon, an anonymous reviewer, and the FLoSS reading group. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 655577 (LOVe) and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 715154, AMORE); Spanish MINECO grants FFI2010-15006, FFI2010-09464-E, and FFI2013-41301, as well as an ICREA Academia award to the first author. This paper reflects the authors' view only, and the funding agencies are not responsible for any use that may be made of the information it contains.

References

- Asher, N. (2011). *Lexical meaning in context*. Cambridge: Cambridge University Press.
- Bally, C. (1944). *Linguistique générale et linguistique française*. Berne: A. Francke.
- Baroni, M., Bernardi, R., & Zamparelli, R. (2014a). Frege in space. *Linguistic Issues in Language Technology*, 9(6), 5–110.
- Baroni, M., Dinu, G., & Kruszewski, G. (2014b). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 238–247), Baltimore, Maryland: Association for Computational Linguistics.
- Baroni, M., & Zamparelli, R. (2010). Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of EMNLP* (pp. 1183–1193). Boston, MA.
- Barsalou, L. W. (2017). Cognitively plausible theories of concept composition. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 9–30). Berlin: Springer.
- Bernardi, R., Boleda, G., Fernandez, R., & Paperno, D. (2015). Distributional semantics in use. In *Proceedings of EMNLP 2015 Workshop LSDSem 2015: Linking Models of Lexical, Sentential and Discourse-level Semantics* (pp. 95–101).
- Boleda, G. (2007). *Automatic acquisition of semantic classes for adjectives*. Ph.D. thesis, Universitat Pompeu Fabra.
- Boleda, G., Baroni, M., Pham, N. T., & McNally, L. (2013). Intensionality was only alleged: On adjective-noun composition in distributional semantics. In *Proceedings of IWCS 2013*, Potsdam.

²⁰In the examples we gave in Sect. 4, values are discrete; however, in actual distributional semantic models values are typically continuous as a result of the operations made on the original counts. See the references in that section for details.

- Boleda, G., Evert, S., Gehrke, B., & McNally, L. (2012). Adjectives as saturators vs. modifiers: Statistical evidence. In M. Aloni, V. Kimmelman, F. Roelofsen, G. W. Sassoon, K. Schulz & M. Westera (Eds.), *Logic, Language and Meaning—18th Amsterdam Colloquium, Amsterdam, The Netherlands, December 19–21, 2011, Revised Selected Papers*, Lecture Notes in Computer Science 7218 (pp. 112–121). Springer.
- Boleda, G., & Herbelot, A. (2016). Formal distributional semantics: Introduction to the special issue. *Computational Linguistics*, 42(4), 619–635.
- Bosch, P. (1983). “Vagueness” is context-dependence. A solution to the Sorites Paradox. In T. Ballmer & M. Pinkal (Eds.), *Approaching Vagueness* (pp. 189–210). North Holland, Amsterdam.
- Bouchard, D. (2005). Sériation des adjectifs dans le SN et formation de concepts. *Recherches Linguistiques de Vincennes*, 34, 125–142.
- Burnage, G., & Dunlop, D. (1992). Encoding the British National Corpus. In J. Aarts, P. de Haan & N. Oostdijk (Eds.), *English Language Corpora: Design, Analysis and Exploitation. Papers from the Thirteenth International Conference on English Language Research on Computerized Corpora* (pp. 79–95). Rodopi, Amsterdam.
- Byrne, D., & Kölbel, M. (Eds.). (2009). *Arguing about Language*. London: Routledge.
- Carlson, G. N. (1977). *Reference to kinds in English*. Ph.D. thesis, University of Massachusetts at Amherst.
- Chatzkyriakidis, S., & Luo, Z. (Eds.). (2017). *Modern perspectives in type-theoretical semantics*. Berlin: Springer.
- Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*, 15, 181–195.
- Clark, H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.
- Coecke, B., Sadzadeh, M., & Clark, S. (2011). Mathematical foundations for a compositional distributed model of meaning. *Linguistic Analysis*, 36, 345–384.
- Del Pinal, G. (2015). Dual content semantics, privative adjectives, and dynamic compositionality. *Semantics and Pragmatics*, 8(7), 1–53.
- Erk, K. (2012). Vector space models of word meaning and phrase meaning: A survey. *Language and Linguistics Compass*, 6(10), 635–653.
- Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25–50.
- Gagné, C. L., Spalding, T. L., & Kostelecky, M. (2017). Conceptual combination, property inclusion, and the Aristotelian-Thomistic view of concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 223–244). Berlin: Springer.
- Garrette, D., Erk, K., & Mooney, R. (2011). Integrating logical representations with probabilistic information using Markov logic. In *Proceedings of IWCS 2011* (pp. 105–114).
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Hampton, J. A. (2017). Compositionality and concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 31–94). Berlin: Springer.
- Jespersen, O. (1942). *A modern English grammar on historical principles*. Copenhagen: Einar Munksgaard.
- Kamp, H. (1975). Two theories about adjectives. In E. Keenan (Ed.), *Formal semantics of natural language* (pp. 123–155). Cambridge: Cambridge University Press.
- Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen & M. Stokhof (Eds.), *Formal Methods in the Study of Language* (Vol. 1, pp. 277–322). Amsterdam: Mathematisch Centrum.
- Kamp, H., & Reyle, U. (1993). *From discourse to logic*. Dordrecht: Kluwer.
- Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–614). New York: Oxford University Press.
- Kayne, R. (1984). *Connectedness and binary branching*. Dordrecht: Foris.
- Kennedy, C., & McNally, L. (2010). Color, context, and compositionality. *Synthese*, 174, 79–98.

- Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211–240.
- Landman, M. (2001). *Adjectival modification restricted*. Ms: University of Massachusetts Amherst.
- Larson, R. (1998). Events and modification in nominals. In D. Strohovitch & A. Lawson (Eds.), *Proceedings from SALT VIII* (pp. 145–168). Ithaca, NY: CLC Publication.
- Lenci, A. (2008). Distributional approaches in linguistic and cognitive research. *Italian Journal of Linguistics*, 20, 1–31.
- Levi, J. (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.
- Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretic approach. In R. Bäuerle, C. Schwarze & A. von Stechow (Eds.), *Meaning, use, and interpretation* (pp. 302–323). Berlin: de Gruyter.
- Martin-Löf, P. (1984). *Intuitionistic type theory: Notes by Giovanni Sambin of a series of lectures given in Padua in June 1980*. Studies in proof theory. Bibliopolis, Napoli.
- McNally, L. (2017). Kinds, descriptions of kinds, concepts, and distributions. In K. Balogh & W. Petersen (Eds.), *Bridging formal and conceptual semantics: Selected papers of BRIDGE-14* (pp. 37–59). DUP, Düsseldorf.
- McNally, L. & Boleda, G. (2004). Relational adjectives as properties of kinds. In O. Bonami & P. Cabredo Hofherr (Eds.), *Empirical issues in syntax and semantics*, (vol. 5, pp. 179–196). <http://www.cssp.cnrs.fr/eiss5>.
- Mitchell, J., & Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science*, 34, 1388–1429.
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge: MIT Press.
- Ogden, C. K., & Richards, I. A. (1923). *The meaning of meaning*. London: Kegan Paul.
- Partee, B. H. (2010). Privative adjectives: Subsective plus coercion. In R. Bäuerle, U. Reyle & T. E. Zimmermann (Eds.), *Presuppositions and discourse: Essays offered to Hans Kamp* (pp. 273–285). Bingley, UK: Emerald Group Publishing.
- Pelletier, F. (2017). Compositionality and concepts—A perspective from formal semantics and philosophy of language. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 31–94). Berlin: Springer.
- Pham, N. T. (2016). *Sentential representations in distributional semantics*. Ph.D. thesis, University of Trento.
- Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
- Rietveld, E. (2008). *Unreflective action. A philosophical contribution to integrative neuroscience*. Ph.D. thesis, University of Amsterdam.
- Rothschild, D., & Segal, G. (2009). Indexical predicates. *Mind and Language*, 24(467–493).
- Sadrzadeh, M., & Purver, M. (2015). Geometry of meaning from words to dialogue acts. In *Proceedings of the IWCS 2015 Workshop on Advances in Distributional Semantics*. London, UK.
- Siegel, M. E. (1976). *Capturing the adjective*. Ph.D. thesis, University of Massachusetts at Amherst.
- Socher, R., Huval, B., Manning, C. D., & Ng, A. Y. (2012). Semantic compositionality through recursive matrix-vector spaces. In *Proceedings of the Joint Meeting of the Conference on Empirical Methods in Natural Language Processing and the Conference on Computational Natural Language Learning (EMNLP-CoNLL)*, Jeju Island, Korea. <http://aclweb.org/anthology//D/D12/D12-1110.pdf>.
- Spencer, A. (1999). Transpositions and argument structure. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1998* (pp. 73–101). Dordrecht: Kluwer.
- Turney, P., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37, 141–188.
- Vecchi, E. M., Baroni, M., & Zamparelli, R. (2011). (Linear) maps of the impossible: Capturing semantic anomalies in distributional space. In *Proceedings of the Workshop on Distributional Semantics and Compositionality* (pp. 1–9), Portland, Oregon, USA: Association for Computational Linguistics.

- Westerlund, M., & Pykkänen, L. (2017). How does the left anterior temporal lobe contribute to conceptual combination? Interdisciplinary perspectives. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 269–290). Berlin: Springer.
- Zamparelli, R. (1995). *Layers in the determiner phrase*. Ph.D. thesis, U. Rochester.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



How Does the Left Anterior Temporal Lobe Contribute to Conceptual Combination? Interdisciplinary Perspectives

Masha Westerlund and Liina Pykkänen

Abstract Within the cognitive neuroscience of language, the left anterior temporal lobe (LATL) is one of the most consistent loci for semantic effects; yet its precise role in semantic processing remains unclear. Here we focus on a literature showing that the LATL is modulated by semantic composition even in the simplest of cases, suggesting that it has a central role in the construction of complex meaning. We show that while the LATL's combinatory contribution generalizes across several linguistic factors, such as composition type and word order, it is also robustly modulated by conceptual factors such as the specificity of the composing words. These findings suggest that the LATL's role in composition is at the conceptual as opposed to the syntactic or logico-semantic level, making formal semantic theories of composition less obviously useful for guiding future research on the LATL. For an alternative theoretical foundation, this chapter seeks to connect LATL research to psychological models of conceptual combination, which potentially offer a more fruitful space of hypotheses to constrain our understanding of the computations housed in the LATL. We conclude that, though currently available data on the LATL do not rule out relation-based models, they are most consistent with schema-based models of conceptual combination, with the LATL potentially representing the site of concept schema activation and modification.

This chapter is a lightly revised version of Chap. 4 in Masha Westerlund's 2015 Ph.D. thesis entitled "Unpacking the neural basis of conceptual combination with MEG" (New York University, Department of Psychology).

M. Westerlund (✉)

Department of Psychology, New York University, New York, USA
e-mail: masha.westerlund@gmail.com

L. Pykkänen

Departments of Psychology and Linguistics,
New York University, New York, USA
e-mail: liina.pykkanen@nyu.edu

L. Pykkänen

NYUAD Institute, New York University Abu Dhabi,
Abu Dhabi, United Arab Emirates

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_11

1 Introduction

The ability to build complex concepts from simple parts is fundamental to human language. Thus a central part of the cognitive neuroscience of language is to develop a model of where, when and how this type of composition occurs. Neuroimaging experiments have shown that sentence processing engages a large network of primarily left-lateralized, interconnected regions (Binder et al. 2009; Jefferies 2013). One of these regions in particular, the left anterior temporal lobe (LATL), appears to be directly implicated in the basic process of combining words into phrases, such as the combination of nouns with adjectival modifiers (*red boat*) or verbal predicates (*eats meat*) (Bemis and Pykkänen 2011, 2013a; Pykkänen et al. 2014; Westerlund et al. 2015; Westerlund and Pykkänen 2014). In order to develop and test hypotheses about combinatory responses in the LATL, recent research has tested representational distinctions arising both from linguistic theory and neuropsychological studies of patients with temporal lobe damage. This work has revealed that LATL responses generalize across various word classes (Westerlund et al. 2015), but are sensitive to the conceptual nature of the combining words. For example, the conceptual specificity of the composing items robustly modulates the combinatory responses in the LATL (Westerlund and Pykkänen 2014; Zhang and Pykkänen 2015) and composition such as numeral quantification (e.g., *two boats*), which arguably does not add any conceptual features to a noun, does not elicit LATL effects at all (Del Prato and Pykkänen 2014; Blanco-Elorrieta and Pykkänen 2016). Thus the LATL appears mostly sensitive to the composition of elements whose meanings are in some sense clearly conceptual, as opposed to composition in a more general sense. In other words, the presence of semantic composition alone is not predictive of combinatory effects in the LATL but rather the semantic composition needs to fit a certain a conceptual profile. Given this, the psychological literature on conceptual combination is a potentially a useful cognitive basis for studying the LATL, a connection we explore in this article.

A large psychological literature has carefully investigated the way that meaning is constructed at the conceptual level and developed models of possible combinatory mechanisms (see Murphy 2002 for discussion). Therefore, these models could be useful for generating predictions about the LATL's combinatory role. In this chapter, we outline two major models of conceptual combination, 'schema-based' models and 'relation-based' models and discuss how these models may be able to constrain the hypothesis space regarding the role of the LATL in composition.

2 The LATL as a Central Combinatory Region

The LATL's anatomical location makes it an excellent candidate for a central combinatory region. It is well connected to primary sensory and motor areas, along with their related association cortices (Catani and De Schotten 2008; Gloor 1997),

is close to the medial temporal lobe, which supports memory processes, and is functionally connected to several critical language regions, including the left inferior temporal gyrus and the middle temporal gyrus (Hurley et al. 2014).

Several neuroimaging experiments have implicated the LATL in sentence processing above the word level. It is consistently engaged by contrasts between structured, meaningful sentences and word lists or jabberwocky sentences (Bottini et al. 1994; Brennan et al. 2010; Brennan and Pylkkänen 2012; Crinion et al. 2003; Friederici et al. 2000, 2003; Humphries et al. 2001, 2005, 2006, 2007; Mazoyer et al. 1993; Pallier et al. 2011; Rogalsky and Hickok 2009; Scott et al. 2000; Stowe et al. 1998; Vandenberghe et al. 2002; Xu et al. 2005). Most critically for its putative status as a combinatory region, a series of magnetoencephalography (MEG) experiments has shown the LATL to be the region most reliably engaged by the composition of basic adjective-noun phrases (Bemis and Pylkkänen 2011, 2013b; Del Prato and Pylkkänen 2014; Pylkkänen et al. 2014; Westerlund and Pylkkänen 2014). The LATL shows consistent differences in neural activity, peaking at around 250 ms, between adjective-noun phrases (e.g. *red boat*) and the same nouns in isolation (preceded by an unpronounceable consonant string, e.g. *xqk boat*) (Bemis and Pylkkänen 2011). Importantly, there is no equivalent increase in activity when nouns are presented in a list (e.g. *cup, boat*) rather than in a phrase, suggesting that the LATL is specifically engaged by the composition of words into a coherent phrase, rather than by a simple increase in the amount of lexical information present.

Of course, composition is not limited to the combination of adjectives and nouns, but extends across all word classes: verbs and adverbs, verbs and their objects, prepositions and their arguments, to name but a few. Within linguistics, at least one prominent model of semantic composition outlines two broad types of composition: the satisfaction of a predicate's argument position (*argument saturation*), and the optional modification of a predicate (*modification*) (Heim and Kratzer 1998). Argument saturation represents a core process of semantic composition. Several word classes, such as verbs, prepositions, and determiners, require arguments in order to exist in well-formed expressions. For example, in the phrase 'eat meat', the verb 'eat' takes the direct object 'meat' as its internal argument. Composing the verb with its argument *saturates* its argument requirement, allowing it to be interpreted. While argument saturation is crucial to the construction of well-formed sentences, language also contains optional elements that serve to enrich the meaning of a well-formed expression. For example, though 'I ate meat' is a perfectly interpretable sentence, the meaning conveyed is greatly changed when meat is modified with an adjective such as 'spoiled'. This type of optional composition is typically described as modification.

Westerlund et al. (2015) investigated whether LATL combinatory responses generalize across these two composition types by presenting words in either modification or argument saturation contexts (*black sweater* and *eats meat*). Both types of composition elicited larger responses in the LATL compared to the same second words presented in isolation. In both cases, the effects peaked at ~250 ms. Furthermore, similar results were found in Arabic, a language with post-nominal

adjectives (e.g. *sweater black*). These results provide evidence that the LATL plays a general role in composition, independent of composition type, language, or word order.

The LATL has also been implicated in the combination of simpler concepts to create more complex concepts at the single-word level (e.g. *boy* can be represented by the concepts *young* and *male*) (Baron and Osherson 2011; Baron et al. 2010), which suggests that the LATL's role in composition may be at the conceptual level, rather than at the level of semantic composition as traditionally conceived of in linguistics. Specifically, most linguistic theories of semantic composition would not consider the construction of a complex single word concept out of its primary features to be composition in the same sense as the composition of words into phrases (i.e., the semantic complexity of *boy* would not in most formal semantic theories correspond to structural complexity that required combinatory steps).

Relatedly, neuropsychological investigations of patients with damage to the LATL have shown that their sentence comprehension and production remain mostly intact, particularly if semantic demands are kept low (Cotelli et al. 2007; Gorno-Tempini et al. 2004; Grossman et al. 2005; Hodges et al. 1992; Kapur et al. 1994; Kho et al. 2008; Noppeney et al. 2005; Wilson et al. 2012). Instead, damage to the LATL leads to a disorder called semantic dementia (SD), in which patients suffer a severe, amodal, memory loss for concepts that manifests across a variety of tasks, including picture naming, word-picture matching, delayed-copy picture drawing, and categorization (Gainotti 2006, 2007, 2011; Garrard and Carroll 2006; Garrard and Hodges 2000; Hodges et al. 1992, 1995; Mummery et al. 1999, 2000; Patterson et al. 2006; Rogers et al. 2004; Snowden et al. 1989). One well-documented aspect of the pattern of conceptual memory loss in SD is that more specific concepts are disproportionately affected by LATL degradation. Patients with SD use progressively more general labels and lose the ability to distinguish similar concepts, mistaking, for example, a zebra for a horse.

Westerlund and Pykkänen (2014) investigated the relationship between conceptual specificity and composition effects in the LATL, with the goal of characterizing whether these two variables modulate the same LATL activity. The conceptual specificity of the noun was varied in adjective-noun phrases (e.g. *blue canoe* vs. *blue boat*) with results showing that combinatorial responses in the LATL at ~250 ms were sensitive to the specificity of the noun, with less specific nouns (*blue boat*) eliciting greater combinatory responses in the LATL than more specific nouns (*blue canoe*). For a fuller assessment of the position-by-position interplay between single word specificity and composition, a follow-up experiment manipulated the conceptual specificity of both the modifier and head in noun-noun compounds (e.g. *tomato* vs. *vegetable*, *soup* vs. *dish*) (Zhang and Pykkänen 2015). More specific modifiers (*tomato*) elicited the greatest responses in the LATL when composed with a less specific head (*dish*). In both of these studies, the effects of single word specificity were subtle or even completely absent when composition was not at play, such as in the head word position when no modifier was present (*boat* vs. *canoe*) or in the modifier position when the head word had not yet been seen (*vegetable* _ vs. *tomato* _). In contrast, the effect of specificity on composition

was robust. Thus these data suggest that at least this LATL-localized MEG activity, occurring at ~250 ms, is most strongly driven by the composition of concepts and not by access to already stored representations.

In sum, these combined results show that although LATL composition effects generalize across composition types and word order (Westerlund et al. 2015), they can be significantly diminished if the composition does not result in a substantial boost in the conceptual specificity of the head word: modifying an already specific noun (*blue canoe*; Westerlund and Pykkänen 2014) does not elicit a strong combinatory response perhaps because the noun is already quite specific and modifying a noun with a very general modifier also fails to engage the LATL measurably (*vegetable dish*; Zhang and Pykkänen 2015). Westerlund and Pykkänen (2014) argued that these results suggest that the LATL's role in composition might be directly related to its role in semantic memory.

The pattern of semantic memory loss in the LATL has led researchers to hypothesize that it acts as a 'semantic hub' (Lambon Ralph et al. 2010; Patterson et al. 2007; Rogers et al. 2004; Rogers and McClelland 2004; Rogers and Patterson 2007), and Westerlund and Pykkänen (2014) suggested that the semantic hub model might be extended to account for the LATL's involvement in semantic composition. The semantic hub model assumes that concepts are represented as sets of features (for example, a dog *is brown, has eyes, and barks*), and that these features are represented in a distributed manner across the brain in modality-specific areas (color and shape in visual areas, sound in auditory areas, etc.). However, these distributed features alone are not sufficient to account for the abstract amodal representations that humans have of concepts. Though elephants and mice share very little in common visually, we categorize them both as animals based on other overlapping features. Rogers and colleagues argue that these categorization abilities require the existence of a hub, located in the LATL, which stores amodal representations of concepts by coordinating the distributed features. Because more specific concepts share several overlapping features (e.g. all poodles have curly hair, four legs, and floppy ears), it is more difficult to distinguish their representations, which leaves them more vulnerable to damage as the LATL deteriorates. In other words, when presented with two poodles, as opposed to a poodle and an elephant, the hub must work harder to identify the few distinguishing features of each poodle (say, a few inches difference in size) than to identify one of the many differences between poodles and elephants (color, size, sound, location, etc.). Note that the hub model is most plausibly a model of concrete concepts, unlikely to suffice as a general account of the semantic space which obviously includes concepts and combinations such as *original idea*, lacking any obvious distributed modal features (for discussion, see e.g., Shallice and Cooper 2013).

Proponents of the semantic hub model argue that the pattern of amodal semantic memory loss in SD points to the LATL as the most likely candidate for a semantic hub. However, this model has been criticized on the grounds that the pathology of SD is not always neatly contained in the ATLS (Hodges and Patterson 2007; Brambati et al. 2009), which can make it difficult to assert that the ATLS are the locus of the critical damage (Simmons et al. 2010). Furthermore, in the intact brain,

the LATL is not consistently activated across all tasks that require conceptual processing (Simmons et al. 2010). However, this latter fact may be due in part to technical challenges in imaging the ATLS, related to their proximity to the nasal cavities (see Visser et al. 2010 for discussion). The presence of LATL effects in the MEG experiments listed above, which do not suffer from similar imaging challenges, lend some support to this interpretation.

Critics of the semantic hub model have also pointed to the ATL's engagement in social and emotional processing (Olson et al. 2007; Zahn et al. 2007) and theory of mind tasks (Olson et al. 2007), as well as in the recognition of famous and familiar people (Damasio et al. 2004; Gorno-Tempini et al. 1998; Sergent and Signoret 1992) to argue that the ATL's role is specific to the social domain (Simmons et al. 2010). Yet other competing models of the LATL propose that LATL plays a role in recognizing unique items, such as familiar people and places (Damasio et al. 1996; Grabowski et al. 2001; Gainotti et al. 2003; Gainotti 2007; Ross and Olson 2012); or that that semantic information is simply grouped by semantic categories, organized along an anterior-posterior gradient across the temporal lobes (Chao et al. 1999; Martin and Chao 2001).

It is of course possible that the LATL plays multiple roles in conceptual processing. Though we discuss the LATL as a single unified region for the purposes of simplicity, the LATL in fact has several anatomical and functional subdivisions (Fan et al. 2014; Rogalsky and Hickok 2009; Sanjuán et al. 2014; Visser et al. 2012; Visser and Lambon Ralph 2011). Therefore, it is difficult to determine whether these models represent competing or complementary interpretations of LATL activity. We believe the semantic hub model to be the simplest model that can reconcile results from SD with results from the sentence processing literature. In fact, when a noun is modified by an adjective, the outcome is a modified conceptual representation, and Westerlund and Pykkänen (2014) suggested that the semantic hub might also be involved in the construction of that modified representation. This might then account for the observed interaction within the LATL between conceptual specificity and combinatory responses. However, this can only be a rudimentary starting point for characterizing the precise contribution of the LATL, as the semantic hub theories offer no mechanistic model of how complex concepts are composed. Since the richest body of work on this topic can be found under the term “conceptual combination” within cognitive psychology, we turn next to these models in order to evaluate extant LATL data in light of predictions arising from this work.

3 Theories of Conceptual Combination

Within psychology, theories of conceptual combination have focused almost exclusively on the modification of nouns, either by adjectives (e.g. *red apple*) or by other nouns (e.g. *city dog*). In modification, the concept being modified is the ‘head’, while the other is the ‘modifier’. Psychological models for how concepts are

combined fall into two general camps: ‘schema-based’ models and ‘relation-based’ models. Schema-based models propose that interpretation arises from the features of the individual constituents (Cohen and Murphy 1984; Hampton 1987, 1997; Murphy 1988; Rumelhart 1980; Smith et al. 1988; Wisniewski 1997), whereas relation-based models focus on how interpretation arises from the general relationship between constituents (Downing 1977; Gagné 2001; Gagné and Shoben 1997, 2002; Gleitman and Gleitman 1970; Levi 1978).

3.1 Schema-Based Models

Schema-based models start with the assumption that concepts contain sets of features organized into a ‘schema’ (Cohen and Murphy 1984; Rumelhart 1980). Rather than a disorganized list of features (*is red, is sweet, is round*, etc.), features within a schema are organized into a set of dimensions that are important to that concept. For example, the concept *apple* might have dimensions for *color, shape, and taste*, amongst others, but not, say, for *speed*. Within each dimension, features are weighted according to how typical they are for the concept. In the *color* dimension, *apple* might have the possible features *green, red* and *brown* as features, and *red* would be weighted higher than *brown*.

Within this framework, conceptual combination involves the modification of the head’s schema by a feature of the modifier. Smith et al. (1988) laid out a prominent version of this model, in which each dimension is weighted by how ‘diagnostic’ it is for a concept; in other words, how useful the dimension is for distinguishing it from similar concepts. For example, the *taste* dimension is useful for distinguishing an apple from other small, round objects, as well as from other fruit. When two concepts are composed, the modifier’s feature is placed into a particular dimension in the head’s schema. In *red apple*, for example, the feature *red* is placed into apple’s *color* dimension. The process of modifying a dimension also makes that dimension more diagnostic for the concept. This captures the intuitive fact that knowing that a concept is *red* is even more important for identifying a *red apple* than an *apple*.

While useful, this model struggles to account for instances of modification in which the modifier is more complex than a simple adjective like *red*. Murphy (1988) provides the adjective *corporate* as an example. The dimension that *corporate* modifies in a composed phrase is dependent on the head noun: a *corporate car* is one owned by the company (modifying the *ownership* dimension), a *corporate lawyer* is one who works for the company (modifying the *employment* dimension), and *corporate stationery* has the company’s logo on it (modifying a *visual* dimension). Furthermore, there can be complex interactions between the various dimensions of a concept when it is modified (Medin and Shoben 1988; Murphy 2002)—for example, the modifier *brown* does not just modify the color dimension of *apple*, but also affects our idea of its taste (unpleasant) and texture (mushy).

Murphy (1988, 1990) proposed that these complications can be accounted for by adding world knowledge to the model. Comprehenders can then draw upon their existing knowledge of cars, lawyers and stationery when deciding what dimension is likely to be modified by *corporate*. Furthermore, after the appropriate dimension has been identified and modified, world knowledge can be used to make further inferences about the concept (Murphy 1990). For example, a comprehender might realize that a *brown apple* is probably rotten, and use this knowledge to draw conclusions about its likely taste and smell. One possible mechanism for this process is ‘extensional feedback’ (Hampton 1987)—essentially, particular instances of a composed concept can be retrieved from memory, and those memories can be used to refine the representation of the concept. For example, after reading *brown apple*, you might remember the last brown apple you had the misfortune of biting into, and you can incorporate aspects of that memory (the bad smell, the soft mush, etc.) into your representation of the concept.

3.2 *Relation-Based Models*

The central insight behind relation-based models of conceptual combination is the fact that concepts are often combined according to certain patterns. In the phrases *glass bottle*, *pea soup*, and *leather purse*, the modifiers and heads are in a similar relationship; in all three cases, the head is ‘made of’ the modifier. Relation-based models propose that comprehenders make use of these statistical regularities in their language to constrain the process of composition. For example, when a comprehender reads the phrase *glass bottle*, she can recognize an instance of ‘made of’ relationship, and therefore immediately understand that a glass bottle is a bottle ‘made of’ glass without needing to access the specific features of the particular concepts.

The most prominent relational model of conceptual combination is the RICE (Relational Interpretation Competitive Evaluation) model (Spalding et al. 2010), previously the CARIN (Competition Among Relations in Nominals) model (Gagné and Shoben 1997). This model argues that there is a fixed set of ‘primitive’ relations into which all combinations can be classified (Downing 1977; Levi 1978), and that comprehenders store distributional information about the types of relations in which the words of their language tend to occur. The process of combining concepts therefore consists of retrieving the appropriate relationship between the modifier and head and linking the concepts according to that relation. As an example, the noun *mountain*, when used as a modifier, occurs much more often in a *location* relation (e.g. *mountain cloud*) than in an *about* relation (e.g. *mountain magazine*). Gagné and Shoben (1997) showed that participants were quicker at judging the sensibility of phrases in which modifiers were in a more frequent relation to the head (i.e. *mountain cloud* was easier to interpret than *mountain magazine*).

Several researchers have argued that the assumption that there is a constrained set of relational ‘primitives’ that are stored separately in the mental lexicon is not

necessary for a relation-based model. Using Wordnet (Seco et al. 2004), Devereux and Costello (2005) examined the semantic similarity of the constituent words for all the phrases used in Gagné (2001), and concluded that semantically similar words tend to be used in similar kinds of relations. For instance, *gas* and *propane* are lexically very similar, and also tended to be used in combinations with similar kinds of relations (*gas crisis*, *propane shortage*). Maguire et al. (2010) confirmed that the type of relations in which constituents tend to occur can be predicted by the semantic nature of the constituents, in a study of all noun-noun combinations in the British National Corpus. Using WordNet, they classified all the constituents in the corpus into 25 different high-level semantic categories, such as ‘substance’, ‘artifact’, or ‘emotion’, and found a strong relationship between the semantic categories of the constituents in the combinations and their relations. They argued that distributional information based on the semantics of the constituents can guide the construction of a relation between them without necessitating the assumption of a separately stored relation.

Importantly, there is evidence that people are sensitive to these category-level statistical patterns. Maguire et al. (2010) showed that participants were able to evaluate a combination like *chocolate taste* just as quickly as a combination like *chocolate rabbit*, even though only *chocolate rabbit* conforms to the highest relation frequency for chocolate (‘made of’). The RICE model would instead have predicted that *chocolate rabbit* would be interpreted more rapidly, as it is the phrase with the more frequent relation for *chocolate*. The authors suggested that this contradictory result arose from the fact that substance-attribute combinations like *chocolate taste* most frequently occur in the ‘has’ relation, which can make up for the specific preference that chocolate has for a ‘made of’ relation like *chocolate rabbit*.

3.3 Summary

In sum, schema-based models and relation-based models each address different aspects of the composition process. Schema-based models focus on the internal conceptual structure of constituents, and on how this internal structure is modified when the concepts are combined. Relation-based models focus on statistical regularities in how concepts tend to compose, and argue that this information guides the composition process. Importantly, these models are not necessarily mutually exclusive. Maguire et al. (2010) suggest that statistical information could serve a similar function to world knowledge within the schema modification process, by identifying candidate dimensions for modification based on previously encountered similar examples. They provide the example of *stone squirrel*. If a comprehender matches the phrase to the common combination pattern ‘substance-object’, and knows that this combination is most often used in a ‘made of’ relation, he can then avoid retrieving irrelevant features such as *is alive* or *eats nuts*, and instead be guided towards form features like *has a tail* and *has four legs*.

4 Processing Predictions of Schema and Relation-Based Models

4.1 *Storage and Retrieval*

The most important assumption of schema-based models of conceptual combination is that concepts are represented as a collection of features organized into structured schemas. This assumption is also the basis of the semantic hub model of the LATL, in which features are organized in modality specific areas, representing dimensions, and the hub serves to organize all the relevant dimensions into a schema.

Composing concepts requires the selection of the relevant feature of the modifier, as well as the corresponding dimension of the head; therefore, schema-based models also assume that composition first requires the retrieval of the feature representations of both constituent concepts. According to the semantic hub model, these feature representations are mediated in the LATL. Together, these models predict that the LATL is involved in retrieving the feature schema for the modifier and head noun, and that this process is the necessary first stage of conceptual combination.

Relation-based models argue that language users store statistical information about the distribution of concepts in their language across relation types. Maguire et al. (2010) suggest that the relevant distributional information about composed phrases is at the level of general semantic categories, such as plant-plant (e.g. *flower bud*) and substance-substance (e.g. *wax paste*), rather than at the level of individual words. Under this type of account, one might then expect certain neural responses to encode combinations of category-level information. However, as regards the LATL, Westerlund and Pykkänen (2014) and Zhang and Pykkänen (2015) have shown that LATL combinatorial responses are sensitive to the specificity of the nouns being composed, even when the nouns are in the same general semantic category (for example, *blue canoe* and *blue boat* might each be categorized as ‘attribute-object’). Therefore, category-level information alone does not appear to be driving LATL responses, though these results of course do not directly rule out relation-based models. Determining whether the frequency of a given combination affects combinatorial responses in the LATL would be a more direct test of the predictions of relation-based models. Furthermore, as proposed by Maguire et al. (2010), it is possible that statistical information is accessed prior to the composition response measured in the LATL, and used to guide feature selection.

More generally, if distributional information does guide the retrieval of features in the LATL, we would expect different representations to be retrieved for the same word, depending on the type of phrase it is presented in. Indeed, there is behavioral evidence that the features that are retrieved for a concept are context-specific (Barclay et al. 1974; Barsalou 1982; Tabossi and Johnson-Laird 1980). For example, Tabossi and Johnson-Laird (1980) presented the same word within two different contexts that each emphasized a different feature of the concept.

- (1) The goldsmith cut the glass with the diamond
- (2) The mirror dispersed the light from the diamond

Participants were faster at verifying the truth of the subsequent sentences *diamonds are hard* or *diamonds are brilliant* when these were compatible with the feature emphasized by the context. However, there is as of yet no specific evidence that distributional information of the kind posited by Maguire et al. (2010) constrains feature selection in the LATL.

In sum, schema models, which assume that a concept's feature schema is retrieved prior to composition, have direct theoretical overlap with semantic hub models of the LATL. The idea that distributional information might direct feature retrieval is open to further investigation.

4.2 Composition

4.2.1 What Is the Combinatory Process?

After the features and/or relations of the constituent concepts have been retrieved, a composed concept is created. A composed concept is not simply the knowledge that, for example, *cooked pasta* is *pasta that has been cooked*. Above and beyond the features of the constituents, composed concepts have features that are true of the combined representation but not of the individual constituents, or 'emergent features'. For example, comprehenders know that *cooked pasta is soft* even though pasta itself could be hard. Thus composition can also result in the 'deletion' of features. Understanding how, when, and even *whether* comprehenders arrive at a complete composed representation is crucial to understanding the combinatory process.

In schema-based models, the core process underlying combination involves the selection of a relevant feature of the modifier and the related dimension of the head. Thus, these models predict that it will be more difficult to select the appropriate features for constituents with more complex internal representations, and that such phrases will therefore be more difficult to compose. There is behavioral evidence that this is in fact the case. Participants were slower to understand novel phrases with noun or non-predicating adjectives modifiers (such as *prostitute committee*) than novel phrases with predicating adjectives (such as *edible food*) (Murphy 1990). The speed of interpretation was also affected by the typicality and abstractness of the modifier (*inedible food* was interpreted more slowly than *edible food*) (Murphy 1990; van Jaarsveld and Drašković 2003; Xu and Ran 2011). If the LATL is involved in the process of composing concepts, rather than in simply retrieving the relevant representations, we might expect to see greater LATL activity for more complex modifiers. This is consistent with the results of Zhang and Pykkänen (2015), in which composing more specific modifiers, with more features, elicited greater LATL activity on the head noun.

It is less clear how to reconcile this prediction with results from Westerlund and Pykkänen (2014) showing that *less* specific heads elicit greater LATL combinatory responses. Less specific concepts have fewer features and therefore a less complex internal representation; however, it may be more difficult to select the appropriate dimension to modify in a more general concept. For instance, what is a *brown animal*? If it's a mammal, this might be an animal with the feature *brown* in the color-of-fur dimension. Alternatively, it might be color-of-scales if it's a fish, or color-of-feathers if it's a bird. A noun denoting a general category of concepts, comprising multiple disparate subcategories of concepts, may not have a single readily accessible dimension that can be modified, in which case it might be more difficult to compose. Of course, this hypothesis makes the prediction that combinatory responses will vary depending on the modifier; a modifier that targets a feature at the more general level might be easier to compose with a more general head (e.g. *dead animal*).

Schema-based models also assume that a combined concept will inherit most of the features of its constituents. Other than the dimension(s) being modified, the features of the head noun should remain stable in the resulting representation. For example, the representation of a *red apple* should still have the features of *apple* (i.e. *is sweet, is a fruit, grows on a tree, etc.*). Connolly, Fodor, Gleitman, and Gleitman (Gainotti 2007) argue, however, that the feature inheritance assumption is incompatible with a phenomenon they term the 'modification effect'—the fact that properties that are believed to be true of a concept are judged to be less true of a modified concept. As a typical example, people rate *baby ducks have webbed feet* as less true than *ducks have webbed feet*, despite the fact that *webbed feet* should be inherited from *duck* into the composed phrase (Connolly et al. 2007). The modification effect also holds for non-word modifiers, which have no semantic content (e.g. *brinn bottles are cylindrical*) (Gagné and Spalding, 2011).

Spalding and Gagné (2014) also provide intriguing evidence for a 'reverse modification effect', in which noun properties that are evaluated as false (e.g. *candles have teeth*) are judged as less false when the noun is modified (e.g. *purple candles have teeth*). Participants appear to be both less comfortable attributing a true property (having webbed feet) and more comfortable attributing false properties (having teeth) to modified nouns.

Gagné and Spalding (2014) argue that the modification and reverse modification effects are more compatible with relation-based models than schema-based models. In the RICE model, they propose that the initial output of the composition process is an underspecified representation, which they term a 'relational gist', limited to the relation between the two constituent concepts. A comprehender's first-pass understanding of a composed concept might thus correspond roughly to the idea that 'a *mountain cloud* is a cloud *located* in the mountains', or that 'a *corporate car* is a car *used by* a corporation' (Gagné and Spalding 2014; Spalding and Gagné 2014). In other words, they argue that no features are inherited during the first stage of the composition process. Any further information about the composed concept, such as the fact that mountain clouds are white, fluffy, and block out the sun, is accessed if and only if the context necessitates further interpretation.

The intuition that combined concepts are initially underspecified, and are only fully fleshed out if the context requires it is similar to a prominent model of sentence processing in psycholinguistics, in which comprehenders construct representations that are ‘good enough’ for the task of understanding the meaning of the sentence but do not reflect a complete syntactic and semantic analysis of the sentence (Ferreira et al. 2001, 2002; Ferreira and Patson 2007). However, this idea is not as incompatible with schema-based models as Spalding and Gagné suggest. It is somewhat implausible to assume that *every* feature of a concept is retrieved and active during the process of composition: in composing the constituents *baby* and *duck*, for example, a comprehender would have to activate every single feature that he or she knows about ducks, including the fact that they have eyes, a heart and lungs, lay eggs, breathe, etc. Instead, context and world knowledge can limit feature retrieval. Therefore, if one assumes that in the absence of a supporting context the specific feature *having webbed feet* is not part of the initial combined concept of *baby duck*, schema modification models do not strictly conflict with modification and reverse modification effects. Instead, these effects might arise out of a post hoc pragmatic reasoning process which assumes that a speaker is providing exactly the necessary information and no more, leading participants to assume that the modifier plays an important contrastive role in the phrase (Gagné and Spalding 2014; Grice 1975; Hampton et al. 2011; Jönsson and Hampton 2012).

In sum, modification and reverse modification effects themselves are potentially compatible with either model of composition. Instead, the major difference between both models lies in whether the activation of conceptual features is *necessary* in order for composition to occur. Schema-based models posit that conceptual features are the elements being composed, whereas relation-based models assume that a ‘good-enough’ relational gist, lacking specific featural information, is the initial output of the combinatory process.

The fact that the complexity of a modifier’s schema affects composition in the LATL is therefore most compatible with the predictions of schema models. In the absence of a context requiring the retrieval of specific features, the relational gist hypothesis does not straightforwardly predict that the conceptual specificity of the constituents would affect LATL responses. Schema-based models further predict that combinatory processes will be disrupted by damage to the LATL; for example, it should be very difficult to comprehend a phrase in which the modifier addresses a dimension of a more specific concept, since that concept’s representation will be degraded. On the other hand, if distributional information can be used to construct a relational gist without accessing a concept’s features, patients with SD should not struggle to comprehend composed phrases. In this case, we might expect these ‘gist’ representations to reflect category-level generalizations, with patients constructing similar interpretations of, for example, *red apple* and *red cherry*. To the best of our knowledge, patient studies have not yet provided evidence about patients’ combinatory abilities that is fine-grained enough to allow us to verify these predictions.

4.2.2 Timing of Composition

Both schema-based and relation-based models are compatible with the existence of a first-pass, shallow composition process, followed by a more elaborative stage. According to schema-based models, the first stage of the combinatory process involves the retrieval of concept features followed by an elaboration stage aided by world knowledge; therefore, combinatory activity in the LATL, peaking at 250 ms, could provide an estimate of the timing of this first stage of feature retrieval. On the other hand, the first stage of composition in the RICE model is the construction of a relational gist, and concept features are only accessed, if necessary, in a second stage. Therefore, the most straightforward way to reconcile this model with current evidence about the LATL's combinatory responses is to either assume that combinatory activity in the LATL represents a second processing stage after a relational gist has been constructed, or that LATL activity is tangential to the composition process.

Furthermore, because the first stage of composition in the RICE model does not involve the retrieval of conceptual features, this makes the prediction that features that are irrelevant to the combined concept will never be activated at any stage of processing. Several behavioral experiments have been conducted to address this prediction by investigating the accessibility of emergent and deleted features at various points during composition. Early evidence suggested that phrasal features were actually accessible *faster* or at the same time as the features of the individual constituents (Gagné and Murphy 1996; Glucksberg and Estes 2000; Hampton and Springer 1989; Potter and Faulconer 1979). For example, Springer and Murphy (1992) found that subjects were faster at evaluating the truth of a sentence like *boiled celery is crispy* than *boiled celery is green*. At first glance, these results are compatible with the predictions of the RICE model.

However, these early experiments measured relatively slow response times (around one second), and therefore may have primarily reflected late-stage reasoning processes. When McElree et al. (2006) used a speed-accuracy tradeoff task to examine the online accessibility of noun and phrasal features, they showed that noun features (*water pistols have triggers*) were verified more accurately than emergent features (*water pistols are harmless*) at early stages of processing. Importantly, subjects were also slower to reject deleted features (*water pistols are dangerous*), suggesting that irrelevant noun features are retrieved during the combinatory process. Furthermore, Swinney et al. (2007) used a cross-modal priming task to confirm that deleted features (*peeled banana-yellow*) are primed during composition, and more quickly than emergent features (*peeled banana-white*).

These results suggest that irrelevant noun features (e.g. *bananas are yellow*) are in fact activated during composition, contradicting the predictions of relation-based models. Furthermore, phrasal features (e.g. *peeled bananas are white*) emerged relatively rapidly, in the absence of explicit pragmatic demands. This suggests that at least some phrasal features are an automatic outcome of the composition process, though they emerge after noun features have been retrieved.

In sum, behavioral and neurophysiological evidence suggest that constituent features are retrieved very rapidly, at or before 250 ms after the onset of composing word. This is followed by the retrieval of features of the composed phrase, a process that may still be ongoing at around 600 ms (Molinaro et al. 2012). Therefore, current evidence supports a two-stage model of conceptual combination. First, at least some of the features of both constituents are activated. This feature activation may be guided by an existing context, and might also be guided by statistical information about the semantic category of the constituents. Then, the modifier's feature modifies the representation of the head concept, leading to the emergence of new phrasal features and the suppression of irrelevant noun features. This can be followed by a further stage of explicit reasoning processes, possibly only to the extent that this is necessary in the comprehension context (Hampton 1987; Murphy 1988, 1990). Of course, composed phrases are often encountered in the context of a phrase or paragraph, in which case contextual information may be available prior to the start of the composition process.

Though further investigation is necessary to determine precisely how phrasal features are retrieved, and how contextual and world knowledge information guide retrieval, the fact that irrelevant features are retrieved during the composition process is more compatible with schema-based models than with relation-based models, though it remains unclear whether these irrelevant features are retrieved in the LATL.

5 Conclusions and Future Directions

In this chapter, we have laid out the general predictions and assumptions of schema and relation-based models of conceptual combination. Schema-based models argue that composition involves the activation of a both constituents' feature schemata and the modification of the head's schema by a feature of the modifier, whereas relation-based models focus on the importance of distributional information in guiding the interpretation of a composed concept.

The fact that LATL combinatory responses take place mostly on the head, after both feature representations can be accessed, and are sensitive to the conceptual specificity of the composing constituents provides preliminary evidence more compatible with the predictions of schema-based models than with the predictions of relation-based models. Schema-based models can be combined with the semantic hub model to predict that the LATL is the locus of feature retrieval and of schema modification, though it is likely that this latter process involves an interplay between the LATL and other higher-order regions, such as the LIFG (Thompson-Schill et al. 1997, 1999) and possibly the angular gyrus (Molinaro et al. 2015; Price et al. 2015), particularly if world knowledge and contextual information can be used to guide the composition process.

Furthermore, schema-based models make the following testable predictions about the LATL as the neural center of composition: (i) that the specific features

activated for the composed constituents should vary to some degree depending on the surrounding context and possibly on distributional information about the constituents, (ii) that the difficulty of integrating the modifier's feature into the head's schema will affect the amplitude or timing of combinatory responses, (iii) that supporting contextual information should mitigate these effects, and (iv) that phrasal features will be retrieved later in composition. Equipped with these predictions, we can now guide our investigation of the LATL in order to construct a detailed model of its combinatory role and its relationships to other language regions. Of course, we do not have enough evidence to rule out relation-based models entirely, and therefore should further investigate the role of distributional information on LATL responses.

In this review, we have focused exclusively on theories of modification, restricted to the composition of adjective-noun and noun-noun phrases. In light of evidence that the LATL shows similar combinatory responses across composition types (Westerlund et al. 2015), it is important to determine whether we can extend current psychological models of composition to include other composition types.

While this challenge is beyond the scope of this chapter, we do note that Bornkessel-Schlesewsky and Schlewsky (2013) have advanced a general compositional model in which all words have an 'actor-event schema', which focuses on actors and actions. For example, an actor-event schema for *paint* might include the typical actor (humans) as well as the typical actions performed around it. This idea is similar in intent to an idea put forth by Wisniewski (1997) that concept schemas should include 'scenarios' corresponding to verbs describing actions or events relevant to the concept. These schemata are then composed in much the same manner as modified phrases, with the head's schema being altered by the information it is composed with. For example, in a phrase like *the doctor paints*, the actor dimension would be filled with the subject *doctor*. Bornkessel-Schlesewsky and Schlewsky (2013) propose that these general schemas are combined in the LATL, and therefore provide a possible first step towards investigating the composition of other types of concepts than simply adjectives and nouns.

Acknowledgements The writing of this chapter and the work reviewed in it were supported by the National Science and Engineering Research Council of Canada (M.W.), the Fonds de recherche du Québec- Nature et technologies (M.W.), the National Science Foundation Grants BCS-0545186 and BCS-1221723 (L.P.) and grant G1001 from the NYUAD Institute, New York University Abu Dhabi (L.P.).

References

- Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, 13(4), 471–481.
- Baron, S. G., & Osherson, D. (2011). Evidence for conceptual combination in the left anterior temporal lobe. *NeuroImage*, 55, 1847–1852.

- Baron, S. G., Thompson-Schill, S. L., Weber, M., & Osherson, D. (2010). An early stage of conceptual combination: Superimposition of constituent concepts in left anterolateral temporal lobe. *Cognitive Neuroscience*, *1*(1), 44–51.
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*(1), 82–93.
- Bemis, D. K., & Pykkänen, L. (2011). Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *The Journal of Neuroscience*, *31*(8), 2801.
- Bemis, D. K., & Pykkänen, L. (2013a). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*, *23*(8), 1859–1873. doi:[10.1093/cercor/bhs170](https://doi.org/10.1093/cercor/bhs170).
- Bemis, D. K., & Pykkänen, L. (2013b). Combination across domains: An MEG investigation into the relationship between mathematical, pictorial, and linguistic processing. *Frontiers in psychology*, *3*.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. (1460–2199 (Electronic)).
- Blanco-Elorrieta, E., & Pykkänen, L. (2016). Composition of complex numbers: Delineating the computational role of the left anterior temporal lobe. *NeuroImage*, *124*, 194–203.
- Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: A new dorsal–ventral stream model of sentence comprehension. *Brain and Language*, *125*(1), 60–76.
- Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., et al. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain*, *117*(6), 1241.
- Brambati, S., Rankin, K., Narvid, J., Seeley, W., Dean, D., Rosen, H., et al. (2009). Atrophy progression in semantic dementia with asymmetric temporal involvement: A tensor-based morphometry study. *Neurobiology of Aging*, *30*(1), 103–111.
- Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., & Pykkänen, L. (2010). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and Language*.
- Brennan, J., & Pykkänen, L. (2012). The time-course and spatial distribution of brain activity associated with sentence processing. *NeuroImage*, *60*(2), 1139–1148. doi:[10.1016/j.neuroimage.2012.01.030](https://doi.org/10.1016/j.neuroimage.2012.01.030).
- Catani, M., & De Schotten, M. T. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, *44*(8), 1105–1132.
- Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *2*(10), 913–919.
- Cohen, B., & Murphy, G. L. (1984). Models of concepts. *Cognitive Science*, *8*(1), 27–58.
- Connolly, A. C., Fodor, J. A., Gleitman, L. R., & Gleitman, H. (2007). Why stereotypes don't even make good defaults. *Cognition*, *103*(1), 1–22.
- Cotelli, M., Borroni, B., Manenti, R., Ginex, V., Calabria, M., Moro, A., et al. (2007). Universal grammar in the frontotemporal dementia spectrum: Evidence of a selective disorder in the corticobasal degeneration syndrome. *Neuropsychologia*, *45*(13), 3015–3023.
- Crinion, J. T., Lambon Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain*, *126*(5), 1193–1201.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380*(6574), 499–505.
- Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition*, *92*(1–2), 179–229. doi:[10.1016/j.cognition.2002.07.001](https://doi.org/10.1016/j.cognition.2002.07.001).

- Del Prado, P., & Pykkänen, L. (2014). MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. *Language Sciences*, 5, 524.
- Devereux, B., & Costello, F. (2005). Investigating the relations used in conceptual combination. *Artificial Intelligence Review*, 24(3), 489–515.
- Downing, P. (1977). On the creation and use of English compound nouns. *Language*, 810–842.
- Fan, L., Wang, J., Zhang, Y., Han, W., Yu, C., & Jiang, T. (2014). Connectivity-based parcellation of the human temporal pole using diffusion tensor imaging. *Cerebral Cortex*, 24, 3365–3378. doi:10.1093/cercor/bht196.
- Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11(1), 11–15.
- Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research*, 30(1), 3–20.
- Ferreira, F., & Patson, N. D. (2007). The ‘good enough’ approach to language comprehension. *Language and Linguistics Compass*, 1(1–2), 71–83.
- Friederici, A. D., Meyer, M., & Von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, 74(2), 289–300.
- Friederici, A. D., Rüschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex*, 13(2), 170–177.
- Gagné, C. L. (2001). Relation and lexical priming during the interpretation of noun–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 236.
- Gagné, C. L., & Murphy, G. L. (1996). Influence of discourse context on feature availability in conceptual combination. *Discourse Processes*, 22(1), 79–101.
- Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(1), 71.
- Gagné, C. L., & Shoben, E. J. (2002). Priming relations in ambiguous noun–noun combinations. *Memory & Cognition*, 30(4), 637–646.
- Gagné, C. L., & Spalding, T. L. (2011). Inferential processing and meta-knowledge as the bases for property inclusion in combined concepts. *Journal of Memory and Language*, 65(2), 176–192.
- Gagné, C. L., & Spalding, T. L. (2014). Subcategorisation, not uncertainty, drives the modification effect.
- Gainotti, G., Barbier, A., & Marra, C. (2003). Slowly progressive defect in recognition of familiar people in a patient with right anterior temporal atrophy. *Brain*, 126(4), 792.
- Gainotti, G. (2006). Anatomical functional and cognitive determinants of semantic memory disorders. *Neuroscience and Biobehavioral Reviews*, 30(5), 577–594.
- Gainotti, G. (2007). Different patterns of famous people recognition disorders in patients with right and left anterior temporal lesions: A systematic review. *Neuropsychologia*, 45(8), 1591–1607.
- Gainotti, G. (2011). The format of conceptual representations disrupted in semantic dementia: A position paper. *Cortex*, 48(5), 521–529.
- Garrard, P., & Carroll, E. (2006). Lost in semantic space: A multi-modal, non-verbal assessment of feature knowledge in semantic dementia. *Brain*, 129(5), 1152–1163.
- Garrard, P., & Hodges, J. (2000). Semantic dementia: Clinical, radiological and pathological perspectives. *Journal of Neurology*, 247(6), 409–422.
- Gleitman, L. R., & Gleitman, H. (1970). *Phrase and paraphrase: Some innovative uses of language*. New York: Norton.
- Gloor, P. (1997). *The temporal lobe and limbic system*. USA: Oxford University Press.
- Glucksberg, S., & Estes, A. (2000). Feature accessibility in conceptual combination: Effects of context-induced relevance. *Psychonomic Bulletin & Review*, 7(3), 510–515.

- Gorno-Tempini, M. L., Rankin, K. P., Woolley, J. D., Rosen, H. J., Phengrasamy, L., & Miller, B. L. (2004). Cognitive and behavioral profile in a case of right anterior temporal lobe neurodegeneration. *Cortex*, *40*(4–5), 631–644.
- Gorno-Tempini, M. L., Price, C., Josephs, O., Vandenberghe, R., Cappa, S., Kapur, N., et al. (1998). The neural systems sustaining face and proper name processing. *Brain*, *121*(11), 2103–2118.
- Grabowski, T. J., Damasio, H., Tranel, D., Ponto, L. L. B., Hichwa, R. D., & Damasio, A. R. (2001). A role for left temporal pole in the retrieval of words for unique entities. *Human Brain Mapping*, *13*(4), 199–212.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3: Speech acts, pp. 41–58). New York: Academic Press.
- Grossman, M., Rhee, J., & Moore, P. (2005). Sentence processing in frontotemporal dementia. *Cortex*, *41*(6), 764–777.
- Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, *15*(1), 55–71.
- Hampton, J. A. (1997). Conceptual combination. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts, and categories* (pp. 133–159). Cambridge, MA: MIT Press.
- Hampton, J. A., Passanisi, A., & Jönsson, M. L. (2011). The modifier effect and property mutability. *Journal of Memory and Language*, *64*(3), 233–248.
- Hampton, J. A., & Springer, K. (1989). *Long speeches are boring: Verifying properties of conjunctive concepts*. Paper presented at the 30th meeting of the Psychonomic Society, Atlanta.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar* (Vol. 13). Wiley-Blackwell.
- Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: Implications for the organization of semantic memory. *Semantic Knowledge and Semantic Representations*, *3*(3–4), 463–495.
- Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Progressive fluent aphasia with temporal lobe atrophy. *Brain*, *115*(6), 1783–1806.
- Hodges, J. R., & Patterson, K. (2007). Semantic dementia: A unique clinicopathological syndrome. *The Lancet Neurology*, *6*(11), 1004–1014.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*(4), 665–679.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: An fMRI study. *NeuroImage*, *36*(3), 924–932. doi:10.1016/j.neuroimage.2007.03.059.
- Humphries, C., Love, T., Swinney, D., & Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping*, *26*(2), 128–138.
- Humphries, C., Willard, K., Buchsbaum, B., & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: An fMRI study. *NeuroReport*, *12*(8), 1749.
- Hurley, R. S., Bonakdarpour, B., Wang, X., & Mesulam, M.-M. (2014). Asymmetric connectivity between the anterior temporal lobe and the language network.
- Jefferies, E. (2013). The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. *Cortex*, *49*(3), 611–625.
- Jönsson, M. L., & Hampton, J. A. (2012). The modifier effect in within-category induction: Default inheritance in complex noun phrases. *Language and Cognitive Processes*, *27*(1), 90–116.
- Kapur, N., Barker, S., Burrows, E., Ellison, D., Brice, J., Illis, L., et al. (1994). Herpes simplex encephalitis: Long term magnetic resonance imaging and neuropsychological profile. *Journal of Neurology, Neurosurgery and Psychiatry*, *57*(11), 1334–1342.
- Kho, K. H., Indefrey, P., Hagoort, P., Van Veelen, C., Van Rijen, P., & Ramsey, N. (2008). Unimpaired sentence comprehension after anterior temporal cortex resection. *Neuropsychologia*, *46*(4), 1170–1178.

- Lambon Ralph, M. A., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences*, 107(6), 2717.
- Levi, J. N. (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.
- Maguire, P., Maguire, R., & Cater, A. W. S. (2010a). The influence of interactional semantic patterns on the interpretation of noun–noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(2), 288.
- Maguire, P., Wisniewski, E. J., & Storms, G. (2010b). A corpus study of semantic patterns in compounding. *Corpus Linguistics and Linguistic Theory*, 6(1), 49–73.
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, 11(2), 194–201.
- Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, 5(4), 467–479.
- McElree, B., Murphy, G. L., & Ochoa, T. (2006). Time course of retrieving conceptual information: A speed-accuracy trade-off study. *Psychonomic Bulletin & Review*, 13(5), 848–853.
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, 20(2), 158–190.
- Molinario, N., Carreiras, M., & Duñabeitia, J. A. (2012). Semantic combinatorial processing of non-anomalous expressions. *NeuroImage*, 59(4), 3488–3501.
- Molinario, N., Paz-Alonso, P. M., Duñabeitia, J. A., & Carreiras, M. (2015). Combinatorial semantics strengthens angular-anterior temporal coupling. *Cortex*.
- Mummery, C., Patterson, K., Price, C., Ashburner, J., Frackowiak, R., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, 47(1), 36–45.
- Mummery, C., Patterson, K., Wise, R., Vandenberg, R., Price, C., & Hodges, J. (1999). Disrupted temporal lobe connections in semantic dementia. *Brain*, 122(1), 61.
- Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science*, 12(4), 529–562.
- Murphy, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language*, 29(3), 259–288.
- Murphy, G. L. (2002). *The big book of concepts*. The MIT Press.
- Noppeney, U., Price, C. J., Duncan, J. S., & Koepp, M. J. (2005). Reading skills after left anterior temporal lobe resection: An fMRI study. *Brain*, 128(6), 1377–1385.
- Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: A review of findings on social and emotional processing. *Brain*, 130(7), 1718–1731.
- Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, 108(6), 2522.
- Patterson, K., Lambon Ralph, M. A., Jefferies, E., Woollams, A., Jones, R., Hodges, J. R., et al. (2006). “Presemantic” cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience*, 18(2), 169–183.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976–987.
- Potter, M. C., & Faulconer, B. A. (1979). Understanding noun phrases. *Journal of Verbal Learning and Verbal Behavior*, 18(5), 509–521.
- Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2015). Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. *The Journal of Neuroscience*, 35(7), 3276–3284.
- Pykkänen, L., Bemis, D. K., & Blanco Elorrieta, E. (2014). Building phrases in language production: An MEG study of simple composition. *Cognition*, 133(2), 371–384.
- Rogalsky, C., & Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebral Cortex*, 19(4), 786.

- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*(1), 205.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. MIT press.
- Rogers, T. T., & Patterson, K. (2007). Object categorization: Reversals and explanations of the basic-level advantage. *Journal of Experimental Psychology: General*, *136*(3), 451.
- Ross, L. A., & Olson, I. R. (2012). What's unique about unique entities? An fMRI investigation of the semantics of famous faces and landmarks. *Cerebral Cortex*, *22*(9), 2005–2015.
- Rumelhart, D. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Erlbaum.
- Sanjuán, A., Hope, T. M., Jones, Ö. P., Prejawa, S., Oberhuber, M., Guerin, J., Price, C. J. (2014). Dissociating the semantic function of two neighbouring subregions in the left lateral anterior temporal lobe. *Neuropsychologia*.
- Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400–2406.
- Seco, N., Veale, T., & Hayes, J. (2004). *An intrinsic information content metric for semantic similarity in WordNet*. Paper presented at the ECAI.
- Shallice, T., & Cooper, R. P. (2013). Is there a semantic system for abstract words? *Frontiers in Human Neuroscience*, *7*, 175. <http://doi.org/10.3389/fnhum.2013.00175>.
- Sergent, J., & Signoret, J. L. (1992). Functional and anatomical decomposition of face processing: Evidence from prosopagnosia and PET study of normal subjects. *Philosophical Transactions of the Royal Society London: Series B, Biological Sciences*, *335*, 55–62.
- Simmons, W. K., Reddish, M., Bellgowan, P. S. F., & Martin, A. (2010). The selectivity and functional connectivity of the anterior temporal lobes. *Cerebral Cortex*, *20*(4), 813–825. doi:10.1093/cercor/bhp149.
- Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, *12*(4), 485–527. doi:10.1207/s15516709cog1204_1.
- Snowden, J. S., Goulding, P., & Neary, D. (1989). Semantic dementia: A form of circumscribed cerebral atrophy. *Behavioural Neurology*.
- Spalding, T. L., & Gagné, C. L. (2014). Property attribution in combined concepts.
- Spalding, T. L., Gagné, C. L., Mullaly, A. C., & Ji, H. (2010). Relation-based interpretation of noun-noun phrases: A new theoretical approach. *Linguistische Berichte Sonderheft*, *17*, 283–315.
- Springer, K., & Murphy, G. L. (1992). Feature availability in conceptual combination. *Psychological Science*, *3*(2), 111–117.
- Stowe, L. A., Broere, C. A. J., Paans, A. M. J., Wijers, A. A., Mulder, G., Vaalburg, W., et al. (1998). Localizing components of a complex task: Sentence processing and working memory. *NeuroReport*, *9*(13), 2995.
- Swinney, D., Love, T., Walenski, M., & Smith, E. E. (2007). Conceptual combination during sentence comprehension evidence for compositional processes. *Psychological Science*, *18*(5), 397–400.
- Tabossi, P., & Johnson-Laird, P. (1980). Linguistic context and the priming of semantic information. *The Quarterly Journal of Experimental Psychology*, *32*(4), 595–603.
- Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron*, *23*(3), 513–522.
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, *94*(26), 14792–14797.
- van Jaarsveld, H., & Drašković, I. (2003). Effects of collocational restrictions in the interpretation of adjective-noun combinations. *Language and Cognitive Processes*, *18*(1), 47–60.
- Vandenberghe, R., Nobre, A., & Price, C. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience*, *14*(4), 550–560.

- Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, 22(6), 1083–1094.
- Visser, M., & Lambon Ralph, M. A. (2011). Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *Journal of Cognitive Neuroscience*, 23(10), 3121–3131.
- Visser, M., Jefferies, E., Embleton, K. V., & Lambon Ralph, M. A. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: Distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *Journal of Cognitive Neuroscience*, 24(8), 1766–1778.
- Westerlund, M., Kastner, I., Al Kaabi, M., & Pykkänen, L. (2015). The LATL as locus of composition: MEG evidence from English and Arabic. *Brain and Language*, 141, 124–134.
- Westerlund, M., & Pykkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia*, 57, 59–70.
- Wilson, S. M., Galantucci, S., Tartaglia, M. C., & Gorno-Tempini, M. L. (2012). The neural basis of syntactic deficits in primary progressive aphasia. *Brain and Language*, 122(3), 190–198.
- Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, 4(2), 167–183.
- Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, 25(3), 1002–1015.
- Xu, X., & Ran, B. (2011). Familiarity, abstractness, and the interpretive strategies of noun-noun combinations. *Cognitive Sciences*, 5, 125–138.
- Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences*, 104(15), 6430–6435.
- Zhang, L., & Pykkänen, L. (2015). The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. *NeuroImage*, 111, 228–240.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Dimension Accessibility as a Predictor of Morphological Gradability

Galit W. Sassoon

Abstract Existing formal theories represent the interpretations of gradable predicates in terms of single scalar dimensions. This paper presents a new approach, which aims to cover morphological gradability in multidimensional adjectives and nouns. Following psychological theories, nouns are assumed to be associated with dimension sets, like adjectives. Degree constructions are proposed to involve quantification on dimensions. This approach correlates the acceptability of a given noun or adjective in comparison constructions with its type of characteristic categorization criterion (i.e., whether, as a default, its dimensions combine into a single criterion via quantifiers or other operations). A preliminary study confirms the predicted correlation. Directions for future research are proposed.

1 Adjectives Versus Nouns in Comparison Constructions

1.1 *The Challenges*

In model-theoretic referential semantics, the interpretation of a predicate (a word like *dog*, *big*, or *card game*) is modeled through its intension, namely, a function from contexts (such as worlds, times, or information states) into classes of entities, contextually given extensions. By contrast, in cognitive psychology, concepts such as ‘dog’, ‘big’, or ‘card game’ are modeled through their dimensions, prototypes, and similarity structures. Gradually, it has been understood that both formal and conceptual representations play a role in natural language semantics. This is evident in the study of morphological gradability, which is the topic of the present study.

Natural language predicates can be divided into various syntactic categories, two of which are adjectives and nouns. Most adjectives, including, *tall*, *expensive* and *healthy* among others, are *morphologically gradable*. In other words, they felicitously

G.W. Sassoon (✉)

Department of English Literature and Linguistics, Bar Ilan University, Ramat Gan, Israel
e-mail: galitadar@gmail.com

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_12

291

combine with degree morphemes, as in *taller*, *tallest*, *too tall*, *tall enough* and *very tall*. However, some adjectives exist that are not morphologically gradable, including, for instance, *geological*, *prime*, and *even*. Thus, a map can be said to be *more expensive* than another map, but not *more geological*.

Degree morphemes that classically combine with gradable adjectives are, by and large, incompatible with nouns, as the infelicitous forms in (1) illustrate. The situation persists across languages (Baker 2003) In particular, within-noun comparisons such as (2) are not naturally used to compare, e.g., real ducks. Their status improves only when, e.g., toys or drawings of ducks are discussed, namely in non-literal interpretations of *duck* ('duck-like'). In English, when the noun occurs as the complement of a mediating particle, such as the preposition *of* in (3a), this also raises the acceptability of the within-noun comparison. Moreover, given nouns like *duck* as inputs, morpho-syntactic processes may yield adjective phrases as outputs, such as *duckish* or *duck-like*, which are compatible with degree morphemes, as illustrated in (3b).

- (1) #*Ducker*, #*Duckest*, #*duck enough*, #*too duck*, #*very duck*
- (2) #The rightmost bird is more a duck than the leftmost bird.
- (3) a. The rightmost bird is more **of** a duck than the leftmost bird.
 b. This toy bird is more {**duckish**, **duck-like**} than that toy bird.

Hence nouns are generally thought to be incompatible with degree morphemes. While the data can be explained by stipulating a syntactic constraint along these lines, researchers of gradability explain the data by postulating a semantic type difference. Nouns are conventionally thought to denote, in each context of evaluation, sets of entities, whereas adjectives are often thought to denote a scalar dimension. For example, in every world, *long* denotes a function, f_{long} , from entities to their degrees on the length dimension. Non gradable adjectives are represented in terms of binary dimensions (Kennedy 1999; see von Stechow 1984 and Heim 2000 for slightly more complex degree-based types). Dimensions form the basis for categorization, therefore entities whose length exceeds a contextual membership norm are classified as *long*.

The problem with the syntactic-category and/or semantic-type approach comes from gradable constructions that ARE acceptable with nouns. The unacceptable construction in (2) above is a *within-predicate comparison*, namely, a comparison of two entities along the dimensions of a single predicate (*duck*). However, the *between-predicate comparisons* in (4a) and (5a, b), which involve comparisons of single entities along the dimensions of two different predicates, seem to have different distributional constraints. Many dimensional adjective pairs cannot naturally co-occur in such comparisons, as illustrated in (4b) (Kennedy 1999), while between-noun comparisons, such as (5a, b), appear generally natural—clearly more than their within-noun counterparts in (5c, d). A mediating particle *of* is mandatory for the latter to sound natural.

- (4) a. This dish is more sour than sweet
 b. #The table is {longer, more long} than heavy
- (5) a. Rubinstein is more a pianist than a conductor.
 b. This creature is more a crab than a lobster.
 c. #Rubinstein is more a pianist than my son
 d. #This creature is more a crab than that one is.

Between-noun comparisons are often analyzed in metalinguistic terms. For example, Morzycki (2011) describes them as comparisons along degrees of imprecision of propositions in the sense of Lasershon (1999). Following this analysis, (5a) conveys that the proposition *Rubinstein is a pianist* is closer to the truth than the proposition *Rubinstein is a conductor*. A problem that arises from this analysis is that both of these propositions are plainly true. Thus, in order to explain examples like (5a), Morzycki's imprecision scale has to include various degrees of plain truth. The basis for the distinction between these degrees seems to lie in the prototypicality structures underlying categorization under *pianist* and *conductor*, but the details of this intuition have yet to be explicated (relevantly, see Hampton 2007 for the view that similarity to a prototype is a common basis for both typicality and graded membership).

In addition, Morzycki (2011) convincingly criticizes previous analyses on the grounds that they are not restrictive enough. However, his account is also not restrictive enough. (5d) is judged less natural than (5b) despite the fact that the propositions *This is a crab* and *That is a crab* may differ in terms of their distance from the truth just as much as the propositions *This is a crab* and *This is a lobster* may. Nor is the data in (1)–(2) explained, e.g., why can't degree morphemes such as *very* and *too* combine with nouns and relate to imprecision-based scales? In sum, metalinguistic scales of imprecision (or of speakers' preference, cf. Giannakidou and Yoon 2011) do not form an exhaustive account of gradability in nouns. This paper proposes that these scales emerge as a consequence of the workings of categorization mechanisms. Only the latter have the potential to explain the complexity of the data.

The goals of this paper are programmatic. It aims to give a direction for an explanation of morphological gradability and to begin developing methods to test its predictions, rather than to report some complete results. The goal is to assess more systematically the relative acceptability of nouns in different comparison types, and to begin testing additional, more specific, predictions of the hypothesized account, while indicating the issues which are left open for future research.

Multiple factors may affect judgments of acceptability, including, for example, syntactic well-formedness, adherence to a language norm or register, frequency of use, and semantic-pragmatic appropriateness (Bard et al. 1996). Focusing on semantic-pragmatic aspects, this paper explores the proposal that the role of their dimensions in categorization is an important factor in explaining the status of predicates in degree constructions. Classification under *long* is merely a matter of length, but classification under *healthy*, for example, is a matter of degree in a

variety of dimensions such as blood-pressure, cholesterol and blood-sugar level. One can be healthy in some respects but not others (Kamp 1975; Klein 1980). Hence, in addition to the fact that, in each context, predicates denote degree functions and are associated with entity sets through classification norms (Kennedy 1999), I propose that the relation *dimension of* associates each predicate (e.g., *bird*, *healthy*, *similar*), in each context, with a set of one or more predicates—its contextual dimensions (e.g., *has a beak*, *healthy with respect to the flu*,).¹

While degree functions and entity sets can be learnt independently of a specification of dimensions, speakers do seem to make pervasive use of dimensions in decisions about degrees and categorization. For example, as explained in Sect. 1.2, the degree functions of nouns are often determined using the degree functions of their dimensions (functions from entities into the degree to which their values in the dimensions match certain ideal values).

Previous formal accounts have considered comparison constructions from the perspective of one-dimensional adjectives such as *long*. However, the naturalness of comparison constructions with multidimensional adjectives, and even with nouns, suggests that the comparison morpheme denotes an operation that can access their dimension set and use it for comparison. An important question arises concerning semantic composition. How do the different dimensions of a predicate and the operation denoted by a degree morpheme combine?

The following paraphrases of hypothesized readings of comparison constructions illustrate the potential relevance of dimensions and the way they combine together. Examples (6a) and (6b) include paraphrases for within-adjective comparisons with a dimensional and a multidimensional adjective, respectively. Examples (6c) and (6d) include possible paraphrases for between-predicate comparisons with multidimensional adjectives and nouns, respectively. A more systematic study is needed to determine whether such readings do indeed exist.

- (6) a. *The sofa is 2 cm longer than the table (is)*: The difference between the degree of the sofa and the table in the dimension underlying entity classification as *long* or not, f_{length} , equals twice the degree of a centimeter.
- b. *(Generally) John is healthier than Bill*: (Generally), the difference between the degrees of John and Bill in the dimensions contextually underlying entity classification as *healthy* or not, $f_{\text{blood pressure}}$, $f_{\text{cholesterol}}$, $f_{\text{chickenpox}}$, ... exceeds zero.
- c. *John is more optimistic than pessimistic*: The number of dimensions along which John is *optimistic* exceeds the number of dimensions along which he is *pessimistic*.

¹Describing in detail a formal model and the ways it addresses philosophical arguments against dimensions would demand more space than allocated for this paper. See Sassoon (2013a) for extensive discussions.

- d. *John is more (of) a linguist than a psychologist*: The number (or proportion) of dimensions of a *linguist* whose membership norm John exceeds is higher than the number (or proportion) of dimensions of a *psychologist* whose norm he exceeds.

These paraphrases suggest that comparison morphemes involve *quantification over* or *counting of dimensions*. In other words, it follows from (6a, b) that a degree difference in at least SOME dimensions of an adjective should be present for within noun comparisons to hold true. In positive multidimensional adjectives, such as *healthy*, the requirement might be stronger to the extent that a degree difference should generally be present in ALL or MOST of the dimensions. Furthermore, it follows from (6c, d) that a larger NUMBER or PERCENTAGE of dimensions of a predicate in comparison to another predicate should be observed in an entity for the comparison to hold true. Without *of*, the comparison in (6d) appears ambiguous between such an interpretation, which seems especially accessible for American English speakers, and an alternative one, whereby the relative similarity of John to the prototypes of *linguist* and *psychologist* compare. Intuitively, John may have neither properties of linguists nor of psychologists, but could still resemble their prototypes to different extents. This would hold true if he is, for instance, a typical philosopher, or maybe a child (Sassoon 2015).

A large body of data suggests that in categorization, the degree functions of the nominal dimensions by default combine to form a single similarity function through *averaging operations* such as weighted sums or products, as assumed in psychological *similarity theories* (cf., Sect. 1.2). By contrast, the degree functions of adjectival dimensions appear by default to combine through *counting operations*, such as those that formal, logical theories assume for quantifying expressions like ALL, MOST or SOME (cf., Sect. 1.3). Thus, the hypothesis presented in (7) emerges:

- (7) a. **The Quantificational Hypothesis**: In within-predicate comparisons (cf., (6a, b), (3)), other degree constructions such as those in (1), and possibly also between-noun comparisons like (5a, b), the degree morpheme combines with the dimensions of the predicates via a counting operation (quantifier) in its semantics.
- b. **The reduced acceptability of nouns**: If the dimensions of a predicate are already bound by similarity operations, it is impossible or hard for a quantifier to access and bind them. As a result of this impossibility (or extra processing cost involved in overriding the similarity operations), acceptability reduces.

The following tests suggest that the dimensions of multidimensional adjectives are more accessible for grammatical operations, and in particular for quantifiers, to operate on, than those of nouns (Bartsch 1986; Landman 1989; Sassoon 2013a). Linguistic labels of operations on dimensions include prepositions, as in (8a), quantifiers, as in (8b), exception phrases, as in (8c), and quantifying adverbs, as in (8d), which on the relevant reading convey that their adjectival argument holds in all (or most) respects. The dimensions of nouns, by contrast, are not easily

accessible in this way, as the reduced acceptability of the nominal examples in (9) illustrates.

- (8) a. healthy **with respect to** blood pressure; good **in** math.
 b. healthy in **{some, most, every}** respect(s); different **in three** respects.
 c. healthy **except for** high cholesterol; identical except in two respects, size and color.
 d. **perfectly** healthy; **{mostly, completely, totally}** different.
- (9) a. #bird with respect to size; #dog in movement.
 b. #tiger in **{some, most, every}** respect(s); #table **in three** respects
 c. #is a bird **except for** size; #is a snake except in two respects, length and color.
 d. **#perfectly** (a) pine; **#{mostly, completely, totally}** a duck.

Accordingly, the within-predicate comparison morpheme (*-er* or *more*) and similar degree morphemes, may denote operations on dimensions which are incompatible with nouns. This is presumably because their dimensions are already bound by similarity operations and thus inaccessible, as stated in (7b). Moreover, according to the quantificational hypothesis presented in (7), the acceptability of within-predicate comparisons (e.g., (6a, b) and (5c, d)) should correlate with the acceptability of constructions involving quantification over dimensions, such as those in (8)–(9). This is because it is hypothesized that the same problem underlies the reduced acceptability of nouns in all cases. The acceptability of these constructions is assumed to depend on the accessibility of the dimensions, namely, the possibility for a quantifying operation to bind them. These predictions extend to similar degree constructions, such as those in (1), and possibly to between-predicate constructions, if those indeed have interpretations such as those in (6c, d). Finally, presumably, mediation by morphemes such as *of* raises acceptability, because these morphemes combine with the noun to prepare a set of dimensions for *more* to operate on. Thus, labor is divided between the two morphemes. However, dimension accessibility may still affect processing and thus acceptability.

Much theoretical work is needed to develop this idea and much experimentation to test it as well. However, this paper focuses on the more basic, general motivation for the quantificational hypothesis. To better understand the determinants of dimension accessibility, the following two sections review studies of the role of dimensions and dimension-binding operations in categorization. We start by discussing the role of similarity operations in categorization under nouns.

1.2 *Psychological Theories of Similarity-Based Categorization*

Consider, for example, ordinary concepts such as *games*. The intuition that all members of such concept categories have something in common has been prevalent

since antiquity. It led to the domination of the classical view of categorization and predicate interpretation as based on definitions, namely necessary and sufficient conditions for membership in the extension. Wittgenstein (1953) encouraged his readers to check whether games such as board games, Olympic games, and word games, have anything in common. This move has led many to reject the classical view on the grounds that definitional properties are rarely, if ever, found. For example, one might think of a game as a physical activity, but this is not valid for chess. One might consider competition against opponents, but solitaire does not involve any opponents. Winning and losing characterize solitaire, but when a child throws a ball at the wall they too disappear. Thus, a member of a natural category may share a slightly different set of properties with each other member. Generally, resemblances between members are much like resemblances between family members.

This conception of categorization in terms of family resemblance triggered the development of fruitful experimental paradigms, including extensive work within cognitive psychology (for a review see Murphy 2002) This work has shown that speakers associate concepts with sets of dimensions, and they systematically consider entities that score highly in these dimensions better examples than others of the concepts in question. For example, some of the *bird* dimensions include *having feathers, wings, beak, small size, egg-laying, flying, singing and perching*. Thus, a robin is considered *more typical* or *representative* of a *bird* than an ostrich (Rosch 1973). When subjects are asked to rate an item by typicality their ratings are relatively similar, and typicality differences between members can be measured with great reliability if one combines the responses of a group of raters (McCloskey and Glucksberg 1978). Furthermore, typicality is a strong predictor of categorization probability (Hampton 1998). and speed (Rosch 1973); e.g., the verification of sentences like *a robin is a bird* is faster than of sentences like *an ostrich is a bird*. But in a context such as *the bird walked across the barnyard*, for which a chicken is regarded as a typical bird, categorization is faster for *chicken* than for *robin* (Roth and Shoben 1983).

These results motivated **the prototype theory**, which models conceptual structure via a set of weighed dimensions and selected dimensional values that characterize what the best example p of that concept is like, whether such an object exists or not. Each dimension F has a *weight* W_F (e.g., W_{flying} tells us how important flying is in classification) and a *selected value*, $f_F(p)$ (e.g., $f_{\text{size}}(p_{\text{bird}})$ represents the ideal size for birds), or range of values (e.g., characteristic colors for apples). The typicality of an item d in a category like *bird* is modeled by d 's similarity to the prototype of bird, p_{bird} ; e.g., the similarity of a robin to a bird is indicated by its averaged degree in the bird dimensions: How well its values match the prototypical values in the dimensions. Since instances of a concept may resemble the prototype in some properties or others, but not necessarily in all its properties, similarity among them is described as family resemblance.

Categorization, following this view, is a process of deciding whether an entity is similar enough to a concept's prototype. Thus, the tight coupling between similarity and membership is captured (Hampton 1998, 2007). Moreover, newly encountered

entities which average better than known members are correctly predicted to be automatically regarded as category members. Thus, this account captures the fact that we are able to categorize infinitely many new instances under the concepts we are familiar with, based on a finite set of encoded dimensions and members. Classification of atypical instances is slower because they have low degrees in the dimensions. Thus, more dimensions have to be considered to determine membership.

Two main types of similarity functions, additive and multiplicative, are relevant to us, because, as we will shortly explain, nouns associated with them are predicted to exhibit different levels of dimension accessibility. *Additive Similarity* can be modeled as the weighted sum of x 's dimensional degrees, as in (10) (Rosch and Mervis 1975),² while *multiplicative similarity* can be modeled by the weighted product of x 's dimensional degrees, as in (11) (Medin and Schaffer 1978).

$$(10) \text{ Additive similarity: } S(x, p) = w_{F_1}f_{F_1}(x) + \dots + w_{F_n}f_{F_n}(x)$$

$$(11) \text{ Multiplicative similarity: } S(x, p) = f_{F_1}(x)^{w_{F_1}} \times \dots \times f_{F_n}(x)^{w_{F_n}}$$

Multiplication models cases in which the most radical decrease in similarity is between entities which perfectly match the ideal in all the dimensions and entities which mismatch in just one or very few dimensions. For instance, assuming, for simplicity, equal dimensional weights and dimensional degrees between 0 and 1, then even instances which match in all of the dimensions except for a 0.5 score in one dimension, have a low mean similarity, 0.5, because multiplication yields $0.5 \times 1 \times \dots \times 1 = 0.5$ (Murphy 2002). Two 0.5 scores yield a mean similarity of 0.25, and so on. The decrease predicted by additive similarity is by far less drastic.

For example, each *bird* sub-species is associated with a bundle of dependent or interrelated dimensions, such as shape, color, behavior, genetic layout, inner biological function and offspring characterization. Multiplication models the fact that a shift in one dimension is usually sufficient to justify classification in a different sub-species; e.g., all robins resemble prototypical robins in being small, eating seeds and fruit, and communicating by singing, whereas all eagles resemble prototypical eagles in being big, eating animals, and communicating with calls. An entity which is somewhat a robin and somewhat an eagle would not classify as either and would, instead, classify as a new species. Thus, multiplicative similarity is recognized by a tendency for interrelated dimensions and gaps between categories, whereas additive similarity is recognized by independent dimensions and category overlaps; e.g., a scholar with some properties of a linguist and some properties of a philosopher may well count as both.

To study dimension integration in different noun domains, Hampton, Storms, Simmons and Heussen (2009) invented scenarios with borderline cases. In one scenario, a historical nuclear accident nearby a remote island resulted in the

²With binary dimensions (dimensions denoting functions from entities to 0 or 1), and weights indicated by the number of category members satisfying them ($w_F = |[[bird]] \cap [[F]]|$), similarity reduces to x 's family-resemblance score—the number of dimensions all-members share with x (cf., Tversky 1977).

evolution of hybrid creatures, e.g., a sub-kind with some features of lobsters and some features of crabs. In a second scenario, a secluded community in a remote area had the habit of using artifacts in ambiguous ways. For instance, some pieces of clothing had features of both a scarf and a tie. Participants were asked to help scientists classify the entities. In line with multiplicative-similarity, hybrids of two natural kinds were often classified as neither one of the kinds. By contrast, in line with additive similarity, hybrids of social categories were often classified as belonging to both categories.

In an additional study, categorization under labels of artifacts and human traits, unlike animals and plants, appeared often to be based on a simple counting strategy, i.e., on whether entities were within the norm in *some* (or *most*) of the concept's dimensions (Wattenmaker 1995). This was modeled with additive similarity, assuming binary dimensions (dimensions which denote functions from entities to 1 or 0) of equal weights. This case is unique in that the effect of averaging can be represented via quantifiers. A social noun *N* is interpreted as denoting the property that an entity *x* has if and only if *x* is within the membership norm ($f_F(x) = 1$) in *SOME* (or *MOST*) of *N*'s dimensions. This property discerns social nouns from natural-kind nouns, in which categorization is based on multiplicative similarity (a weighted *product*, rather than *sum*, as in (11)), so that even a representation of their dimensions as binary and of equal weights does not yield classification based on dimension counting. Thus, given the quantificational hypothesis in (7), social nouns are predicted to be more compatible than natural-kind nouns, not only with quantificational dimension-binding operations, such as those in (8)–(9), but also with degree morphemes such as those in (1)–(6).

Notice, however, that the adjective-noun distinction confounded the results of Wattenmaker's 1995 study. Most of the social concepts in his experiments were adjectives applicable to humans, and most of the natural kind concepts were nouns. Thus, the reported findings indirectly support the view that adjectives and social nouns, but not natural-kind nouns, may be interpreted by means of quantification (dimension counts). The next section reviews linguistic data suggesting that quantification over dimensions is indeed prevalent in categorization under adjectives. But this methodological issue highlights a broader point. To achieve better understanding of the complexity of language cognition, a need arises for tighter collaboration between linguists and psychologists. This need is precisely what this book is meant to address.

1.3 *Logical Operations of Quantification Over Dimensions in Adjective Categorization*

Consider, for example, the adjective *identical*. Native speakers intuitively judge entities as *identical* if and only if they are identical in *all* (or *most*) of the respects that count as relevant in the context. By contrast, they intuitively judge entities to be

different if and only if they are different in *some* (at least *one*) respect. Objects are considered *clean* if and only if dirt of *no sort* (dust, stains, etc.) is attested on them, while they are considered *dirty* if and only if dirt of *some* sort is attested. Individuals with a slight cold are *not* strictly speaking *healthy*, since they are not *all* healthy, while one type of sickness suffices to count as *sick*.

Considering these, among many other examples, the following hypotheses emerge. First, the dimensions of adjectives seem to be integrated using quantification or counting operations. Second, the dimensions of positive adjectives (like, e.g., *identical*, *healthy* and *clean*) appear, as a default, to be integrated by means of *universal quantifiers* (ALL, NO). By contrast, those of their negative antonyms appear to be integrated by *existential quantifiers* (SOME). In other words, intuitively, to count as healthy, one cannot have any serious disease, whereas to count as sick, one must have some disease or other. Hence, *healthy* seems to mean healthy in ALL respects while *sick* means sick in SOME respects. We may consider someone to be healthy (or not to be sick) despite, say, high cholesterol only when this dimension is considered irrelevant. When using expressions like *all* or *everybody*, the standard practice is to ignore irrelevant entities (von Fintel 1994), but not to allow any other exceptions. Moreover, the discourse in a given context may revolve around cholesterol problems, for example, such that *healthy* and *sick* would associate uniquely with this dimension. Thus, positive adjectives also have weak interpretations.

Several corpus and judgment studies (Sassoon 2012, 2013b; Shamir 2013) explored these hypotheses. They suggest that, indeed, universal and existential quantification on dimensions is general among positive and negative adjectives, but rare in concrete nouns. The methodology exploited the fact that exception phrases indicate universal generalizations (generalizations, unlike existence statements, can have exceptions). This fact is illustrated by the higher acceptability of (12a, b) than of (12c, d) (von Fintel 1994; Hoeksema 1995; Moltmann 1995; Fox and Hackl 2006). Notice that, as logical theories predict, negated existential quantifiers are universal (cf., (12b)), and negated universal quantifiers are existential (cf., (12d)).

- (12)
- a. Everyone is happy except for Dan
 - b. No one is happy except for Dan
 - c. #Someone is happy except for Dan
 - d. #Not everyone is happy except for Dan

Accordingly, speakers tend to accept exception phrases with adjectives whose dimensions combine via an implicit universal quantifier ALL, but not with their negative antonyms, whose dimensions combine via an existential quantifier SOME, cf. the acceptability contrast in (13a, c) (Hoeksema 1995). And since negated existential quantifiers are universal and vice versa, the effect is reversed in the presence of negation. Hence, exception phrases tend to be accepted with negated existential adjectives, but not universal ones, as illustrated in (13b, d).

- (13) a. I am **healthy** except for high blood pressure (bp) ($\forall F \neq \text{bp}$, I'm healthy in F).
 b. He is **not sick** except for the flu ($\neg \exists F \neq \text{flu}$, He's sick in F/ $\forall F \neq \text{flu}$, He's healthy in F).
 c. #I am **sick** except for normative blood pressure ($\# \exists F \neq \text{bp}$, \neg (I'm healthy in F)).
 d. #I am **not healthy** except for (normal) cholesterol ($\# \exists F \neq \text{ch}$, \neg (I'm healthy in F)).

These judgments are supported by distributional patterns, as revealed in a study of 1300 naturally occurring examples of the form 'Adj. except' with 8 antonym pairs in positive versus negated contexts (Sassoon 2013b). Frequency of co-occurrence of an adjective with dimensional exception phrases depended both on adjective polarity (positive vs. negative) and context polarity (existence or absence of negation), which interacted significantly ($p < 0.001$). **Universality**, i.e., frequency of implicit universal quantification on dimensions, as measured by the frequency of exception phrase modification of an adjective in contexts without negation, was higher in positive adjectives (cf., (13a)) than in their negative antonyms (cf., (13c)), whereas **existentiality**, i.e. frequency of implicit existential quantification on dimensions, as measured by frequency of exception phrase modification of an adjective when negated, was higher in negative adjectives (cf. (13b)) than in positive ones (cf. (13d)). Moreover, in positive adjectives, universality was higher than existentiality, but in negative adjectives, it was lower.

These patterns were also reflected in survey results (Sassoon 2012). They support the view that negative antonymy systematically affects the force of quantifier on dimensions. The dimensions of positive adjectives tend to be bound by a universal quantifier, ALL or NO, while those of negative antonyms tend to be bound by an existential quantifier, SOME. In other words, by default, multidimensional adjectives are used to convey generalizations on dimensions, while multidimensional antonyms relate to counterexamples to such generalizations. Moreover, Shamir (2013) directly compared adjectival antonyms to concrete nouns, showing that these nouns are less acceptable with dimensional exception phrases than either positive or negative adjectives, in both negated and non-negated contexts ($p < 0.000$; cf. *#bird/not a bird, except for flying/size/wings*), in line with dimension binding via similarity operations, instead of quantification.

At the same time, context can override the tendency for similarity-based dimension binding; e.g., in a scientific context in which birdhood is defined by means of n genes, it is acceptable to state that a certain exemplar *is a bird except with respect to 3 genes*. Similarly, Sassoon's (2013b) corpus results illustrate that context can override the tendency for universal and existential quantification in positive and negative adjectives, respectively. This is also evident from certain intuitive judgments; e.g., both the combination of *smart* and of *stupid* with *except in math* are natural, and so is *The ipad is a bigger iphone except you can't make calls*. In addition, the likelihood of multiplicative versus additive dimension-binding in Wattenmaker's (1995) study was affected by the order of presentation (the number

of preceding additive vs. multiplicative predicates). Other contextual factors await future research (see one pilot in Sassoon 2012).

Returning to default contexts, we have seen that dimension-binding in nouns is mostly based on averaging functions (weighted-sums or -products), rather than quantifiers. This is in accord with the fact that the noun dimensions are normally neither necessary nor sufficient conditions for categorization (Hampton 1995). Noun dimensions are often binary (e.g., ‘wings’ and ‘beak’ for *birds*) and many. In contrast to adjectives, conceptual gradability in nouns emerges due to averaging on many of them, not due to counting or to a choice of a single scalar dimension. Adjectives, by contrast, can be represented as having a free dimension argument F. This argument can be saturated as in *clever in doing math*, which denotes the property of having higher than normal ability to do math, assuming math to be a dimension of *clever*. Alternatively, the dimension argument can be bound by an explicit or implicit quantifying expression as in *vital in every respect* or *atypical in two respects*.

Moreover, the dimension argument can be bound by comparison morphemes; e.g., *more in X is more expensive than Y* selects the unique member of the dimension-set of *expensive* (a function from entities to their cost), and returns the relation of being more costly (that relation between x and y such that for SOME dimension F of *expensive*, x is F-er than y; cf., (6a)). With the multidimensional adjective *healthy*, comparison can be along more dimensions; e.g., *healthier* can denote a relation between entities x and y such that for ALL health dimensions F, x is F-er than y (cf., (6b); for detailed compositional derivations see Sassoon 2014). The hypothesis in (7) proposes that the acceptability of predicates in such constructions depends on the accessibility of their dimensions, namely the extent to which their interpretation can be modeled by means of a dimension argument bound by a quantifier. Here are a number of its testable predictions.

Most importantly, a continuum of (un)acceptability is predicted. All nouns are predicted to be less acceptable than adjectives in within-predicate comparisons, because, by default, their dimensions are bound by averaged similarity operations and thus are not accessible for quantifiers to bind. However, additive nouns are expected to be more acceptable than multiplicative nouns, because under certain assumptions their interpretation can turn to one based on dimension counting. Thus, nouns denoting social categories are expected to be more acceptable than nouns denoting animals and plants (cf., (14c–d) vs. (14a–b), respectively). In addition, if between-noun comparisons exhibit the same continuum (cf., (15c) vs. (15a)), this will support the availability of a reading involving comparison of the number of dimensions of each compared predicate, whose norm an entity exceeds (cf., (6c, d)). But if their only reading directly exploits degrees of similarity of entities to nominal prototypes, no acceptability difference between social and natural-kind nouns is expected.

Similar contrasts are expected to occur also with quantifiers on dimensions and exception phrases that weaken generalizations on dimensions, as in (16) and (8)–(9). Moreover, if comparisons indeed involve quantification on dimensions, then the acceptability of a predicate in comparison constructions and in quantificational constructions, as in (16), is expected to correlate.

- (14) Within-noun comparisons: “X is more NP than that Y”
- #This piece of fruit is more an orange than that piece of fruit.
 - #This farm animal is more a cow than that farm animal.
 - ?This booklet is more a diary than that booklet.
 - ?This artist is more a composer than that artist.
- (15) Between-noun comparisons: “X is more NP₁ than (Y is) NP₂”
- ?This piece of fruit is more an orange than an apple
 - ?This farm animal is more a cow than that farm animal is a horse
 - This booklet is more a diary than a sketchbook
 - This artist is more a composer than that artist is a poet
- (16) Dimensional exception phrases: “X is (not) NP, except for/with respect to DIM”
- #This vegetable is (not) a potato except for color.
 - #This predator is (not) a tiger except for its teeth number.
 - ?This place is (not) a church except in appearance.
 - ?This girl is (not) a genius except with respect to literature.

Finally, within-noun comparisons (cf., (14)) are expected to be less acceptable than single-subject between-noun comparisons (cf., (15a, c)). The latter seem to have a reading that directly exploits the similarity functions underlying categorization in nouns, while the former do not, presumably because they only select for dimensions compatible with difference modifiers (*interval-scale* dimensions), as in *two inches longer* and *slightly shorter* (Sassoon 2013a). Even relatively abstract adjectives seem to allow for an interval-scale construal (e.g., *slightly happier*, *a lot more beautiful*), whereas nouns appear to refuse it and are mostly based on nominal dimensions (e.g., wings: yes/no) or ordinal dimensions (movement type: flying > swimming > running; Gardenfors 2004). Two-subject between-predicate comparisons (cf., (15b, d)) are also expected to dislike nouns as they seem to license difference modifiers (e.g., *The sofa is 2 inches longer than the table is wide*), as opposed to single-subject comparisons (e.g., (15a, c)) which seem to be incompatible with difference modifiers (cf., *#slightly/#a lot more a car than a truck*; Morzycki 2011). The judgment study we now turn to aims to test some of these predictions.

2 A Pilot Study of Acceptability Judgments

2.1 Method

The participants were recruited using Amazon mechanical Turk (AMT), an online labor market place where workers are paid small amounts of money to complete small tasks named HITs (Human Intelligence Tasks). It has been shown that AMT provides a quick and relatively cheap method to acquire high-quality experimental

results that do not differ significantly in performance from standard experimental settings (Buhrmester et al. 2011). The hits were only visible to American workers with approval rates exceeding 95%. They were awarded 2 cents per hit consisting of a single item rating, with an average hourly rate of 6.5\$. A group of 25 different participants completed each hit. Once 25 participants filled out a hit, it was no longer visible. Moreover, each participant chose how many hits to fill out. In total, 140 participants answered an average of 44.5 different questions each (SD = 58). All hits were completed on 18/07/2013.

The *stimuli* were sentences formed of 24 definite noun phrases that served as *subjects* and 24 indefinite noun phrases that served as *predicates*, such as, for example, *This farm animal* and *a cow*, respectively, in *This farm animal is more a cow than that farm animal*. The factors manipulated were *entity type*—half of the subjects (and their predicates) denoted natural entity types (plants and animals) and half denoted social entity type (human traits and artifacts), and *structure*—each subject-predicate combination occurred in 7 different constructions. These constructions were: (i) **two baseline conditions** (Pos, Not) including basic and negated categorization forms ('x is P' and 'x is not P'); (ii) **three comparison conditions** (Within, Be1s, Be2s) including within-noun comparisons ('x is more P than y'; cf., (14)), single-subject between-noun comparisons ('x is more P than Q', as in (15a, c)), and two-subject between noun comparisons ('x is more P than y is Q', as in (15b, d)); and (iii) **two modified forms** (Ex, ExNot) including basic and negated forms with dimensional exception phrases ('x is P except for Dim' and 'x is not P except for Dim', as in (16)).

An eighth condition consisted of within-adjective comparisons, constructed by applying each one of 6 positive one-dimensional adjectival predicates (*tall*, *heavy*, *big*, *old*, *expensive* and *colorful*) to 4 of the 24 subjects, an animal, plant, artifact and human trait. Thus, focusing on the within-predicate comparisons, the factors manipulated were *Predicate type* (adjective vs. noun) and *Entity type* (social vs. natural).^{3,4}

In comparisons with two entities, the two were always described as *This NP* and *That NP* (as in *This artist is more a composer than that artist*.) In comparisons with two predicates, each nominal predicate was matched with a contrasting predicate (e.g., *duck* and *goose*; *dress* and *night shirt*; see Table 1 for the full set of subjects and contrasting predicate pairs). The matched pairs were mostly withdrawn from the stimuli reported in Hampton, Storms, Simmons and Heussen (2009). The pairs of human traits and professions were chosen from a list created by searching the

³An additional factor was *Than-phrase type*—half of the sentences with each entity and predicate type included a clausal *than*-phrase and half included a phrasal one, as in *This animal is taller than that animal is* versus *This person is taller than that person*, respectively.

⁴Here is a set of examples for each of the 8 conditions: (1) This place is a church; (2) This place is not a church; (3) This place is more a church than that place is; (4) This place is more a church than an art gallery; (5) This place is more a church than that place is an art gallery; (6) This place is a church except in appearance; (7) This place is not a church except in appearance; (8) This place is older than that place is.

Table 1 The 24 between noun comparisons with a single subject

Natural entity types			
Animals		Plants	
1	This farm animal is more a cow than a horse	1	This tree is more a pine than an oak
2	This insect is more a mosquito than a wasp	2	This vegetable is more a potato than a carrot
3	This bird is more a crow than a pigeon	3	This flower is more a dandelion than a rose
4	This reptile is more a snake than a lizard	4	This spice is more onion than mint
5	This predator is more a wolf than a tiger	5	This plant is more moss than grass
6	This aquatic mammal is more a dolphin than a shark	6	This piece of fruit is more an orange than an apple
<i>Social entity types</i>			
Human concepts		Artifacts	
1	This journalist is more a consultant than a reporter	1	This place is more a church than an art gallery
2	This artist is more a composer than a poet	2	This piece of clothing is more a dress than a nightshirt
3	This football player is more a champion than a celebrity	3	This car is more a taxi than an ambulance
4	This girl is more a genius than a child	4	This hall is more a theatre than a cinema
5	This colleague is more a businesswoman than a friend	5	This container is more a carafe than a vase
6	This person is more a victim than a witness	6	This booklet is more a diary than a sketchbook

corpus of contemporary American English (Coca, Davies 2010) for the string “(s)he is a” followed by a noun. Unfortunately, two errors occurred in the survey: ‘crew’ and ‘fly’ were used instead of ‘crow’ and ‘insect’, respectively. Thus, the results are reported with these two items removed.

The 192 target sentences (24 × 8 conditions) with 57 additional good and bad fillers of various forms can be found in Appendix 1. The 249 sentences (hits) were presented as a single randomly ordered list, out of which different participants filled out different subsets. The variables are therefore within-subject, but with many missing observations.

The following instructions occurred in every hit: “[This hit is for English Native Speakers Only] RATE THE FOLLOWING SENTENCES BY HOW NATURAL THEY SOUND TO YOU AS AN ENGLISH NATIVE SPEAKER. FOR EXAMPLE: (i) To me, the sentence “This child is rarely sick” is a perfectly natural sentence, so I give it ‘7’; (ii) However, the sentence “This child is rarely tall” is perfectly unnatural and makes no sense (me or my friends would never use such a combination of words), so I give it ‘1’”. Following the instructions the participants saw a sentence and a 7-point scale with the labels PERFECTLY NATURAL and

PERFECTLY UNNATURAL adjacent to 7 and 1, respectively. They selected an answer by clicking on the radio button adjacent to the relevant number.

2.2 Results

The mean naturalness and standard deviation for each condition are presented in Table 2 (for more detailed tables by items see Appendices A-B). The two left columns of Table 2 reveal a potential interaction between *structure* and *entity type*. As predicted, nominal structures involving comparison (conditions 1–3) and exception-phrases (conditions 4–5) were rated as more natural with social items (subjects and predicates) than with natural items (e.g., *more a composer* was rated higher than *more a pine*, and *a diary except for the sketches* was rated higher than *a potato except for color*). However, no such differences were found in adjectival comparisons (condition 6) (e.g., *taller*) and basic categorization structures (7–8) (e.g., *is a pine/composer*).

The right column further suggests that within-adjective comparisons (condition 6), are more natural than single-subject between-noun comparisons (condition 1), which in turn are more natural than within-noun comparisons (condition 2). The least natural seem to be the two-subject between noun comparisons (condition 3).

To test whether these potential effects and interactions are significant, the results of the 7 nominal conditions were analyzed using a variety of mixed models with naturalness as the dependent variable and participant and item (22 subject + predicate combinations) as random effects. This type of analysis tests whether the

Table 2 Averages and standard deviations by condition for 10 natural items (subjects + predicates), 12 social items, and all 22 items

Conditions	M _{natural} (SD)	M _{social} (SD)	M _{total} (SD)
1. Single subject between-noun comparisons (Be1s): More NP1 than NP2	3.94 (2.10)	5.23 (1.90)	4.64 (2.09)
2. Within-noun comparisons (Within): More NP than Y (is)	3.38 (1.98)	4.46 (2.00)	3.97 (2.06)
3. Two-subject between-noun comparisons (Be2s): More NP ¹ than y is NP ²	2.76 (1.74)	3.15 (1.90)	2.97 (1.84)
4. Exception phrases (Ex): X is NP except	2.86 (1.92)	3.51 (2.09)	3.21 (2.04)
5. Negated exception phrases (ExNot): X is not NP except	2.66 (1.80)	3.54 (2.12)	3.14 (2.03)
6. Within-adjective comparisons (Adj): More AP than Y (is)	6.08 (1.58)	6.02 (1.58)	6.05 (1.58)
7. Basic categorization (Pos): X is NP	6.44 (1.28)	6.58 (1.00)	6.52 (1.14)
8. Negated categorization (Not): X is not NP	6.49 (1.24)	6.54 (1.15)	6.52 (1.19)

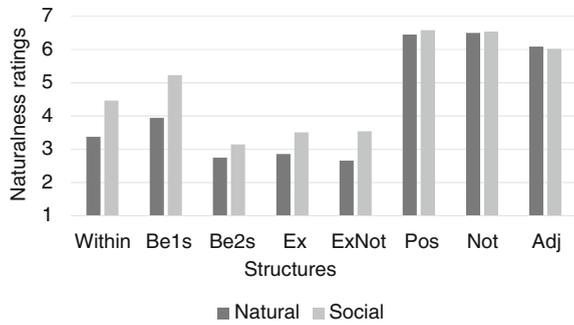
manipulated factors explain anything more than what could be predicted on the basis of the participants and the stimuli. It is reasonably robust against influence of a single outlier item or participant, and against missing observations (Baayen et al. 2006).

The baseline and comparison conditions were compared using a mixed model with Entity type (natural vs. social) and Structure (Basic vs. Comparison) as fixed effects. A structure effect and an interaction were predicted, and indeed a significant main effect was found for both Entity type ($F(1, 20) = 12.2, p = 0.002$) with higher ratings for social ($M = 5.2, SD = 2.1$) than natural items ($M = 4.6, SD = 2.3$), and Structure ($F(1, 2660) = 1979.9, p < 0.001$), with higher ratings for basic ($M = 6.5, SD = 1.2$) than comparative structures ($M = 3.9, SD = 2.1$). Further, the two factors significantly interact ($F(1, 2652) = 45.7, p < 0.001$). Comparative structures were rated higher with social ($M = 4.3, SD = 2.1$) than natural items ($M = 3.4, SD = 2.0$), while basic structures with social ($M = 6.6, SD = 1.1$) and natural items ($M = 6.5, SD = 1.3$) were rated similarly.

Considering the **three comparison conditions** more specifically, a mixed model analysis with structure (Within, Be1s, Be2s) and entity type (social vs. natural) as fixed effects confirmed the predicted entity type effect ($F(1, 20) = 29.3, p < 0.001$), with higher ratings for social ($M = 4.3, SD = 2.1$) than natural items ($M = 3.4, SD = 2.0$), and revealed an additional structure effect ($F(2, 1544) = 140.7, p < 0.001$). Pairwise comparisons confirmed that, as expected, the naturalness of the single-subject between-noun comparisons ($M = 4.6, SD = 2.1$) was significantly higher than the naturalness of the within-noun comparisons ($M = 4, SD = 2.1$; Mean difference 0.65, $SE = 0.096, df = 1544, p < 0.001$), which was significantly higher than the naturalness of the two-subject between-noun comparisons ($M = 3, SD = 1.8$; Mean difference 0.99, $SE = 0.096, df = 1553, p < 0.001$). Univariate tests based on the pairwise comparisons of marginal means yielded $F(2, 1548) = 146.9 (p < 0.001)$.

The analysis further revealed a significant entity by structure interaction ($F(2, 1548) = 13.2, p < 0.001$), with bigger differences between natural and social items in single-subject between-noun comparisons ($M_{\text{social}} = 5.2, SD = 1.9$; $M_{\text{natural}} = 3.9, SD = 2.1$) and within-noun comparisons ($M_{\text{social}} = 4.5, SD = 2$; $M_{\text{natural}} = 3.4, SD = 2$) than in two-subjects between-noun comparisons ($M_{\text{social}} = 3.2, SD = 1.9$; $M_{\text{natural}} = 2.8, SD = 1.7$). To test these **simple entity type effects** (illustrated in Fig. 1), separate mixed models were conducted for each structure, with entity type as the fixed effect. The significantly higher naturalness predicted for social versus natural entity types was confirmed in within noun comparisons ('more N than y'; $F(1, 24) = 13.4, p = 0.001$) and single-subject between-noun comparisons ('more N1 than N2'; $F(1, 28) = 15.4, p < 0.001$), and it approximated significance in two-subject between-noun comparisons ('more N1 than y is N2'; $F(1, 19) = 4.3, p = 0.051$). Importantly, there were no entity type effects in the baselines (for Pos—'x is P': $F(1, 19) = 0.377, p = 0.546$; for Not—'x is not P': $F(1, 19) = 0.046, p = 0.83$) and adjectival condition ('more Adj than y'; $F(1, 20) = 0.122, p = 0.73$). Thus, entity type only affected naturalness of nominal comparisons, as predicted.

Fig. 1 Entity \times Structure interactions: entity type effects (social > natural) occurred in the structures involving nominal comparisons and exception-phrases, while they occurred neither in the basic categorization structures, nor in the adjectival comparison structures



Turning to exception phrases, **the baselines (Pos, Not) and modified conditions (Ex: 'x is N except in dim', ExNot: 'x is not N except in dim')** were also compared using a mixed model with Entity type (natural vs. social) and Structure (Basic vs. modified) as fixed effects. Again, a structure effect and an interaction were predicted. While the Entity type effect was only approximating significance ($F(1, 19) = 4.3, p = 0.051$), with higher ratings for social items ($M = 5, SD = 2.3$) than for natural items ($M = 4.6, SD = 2.4$), the analysis revealed as predicted a significant Structure effect ($F(1, 2109) = 3009.4, p < 0.001$), with higher ratings for basic structures ($M = 6.5, SD = 1.2$) than for exceptive structures ($M = 3.2, SD = 2$), and a significant interaction ($F(1, 2106) = 27.1, p < 0.001$). Exceptive structures were rated higher with social items ($M = 3.5, SD = 2.1$) than with natural items ($M = 2.8, SD = 1.9$), while basic structures with social items ($M = 6.6, SD = 1.1$) and natural items ($M = 6.5, SD = 1.3$) were rated similarly.⁵

Considering the **two exceptive structures** more specifically, a mixed model analysis with structure (Ex, ExNot) and entity type (social vs. natural) as fixed effects confirmed the predicted entity type effect ($F(1, 20) = 5.2, p = 0.033$), with higher ratings for social items ($M = 3.5, SD = 2.1$) than for natural items ($M = 2.8, SD = 1.9$), and with neither a significant structure effect ($M_{Ex} = 3.21, SD = 2; M_{ExNot} = 3.14, SD = 2; F(1, 1008) = 0.818, p = 0.366$) nor an interaction ($F(1, 999) = 2.3, p = 0.131$). Separate mixed models for each condition revealed that, as Fig. 1 illustrates, the structures with exception-phrases exhibited the entity type effect both with negation ($M_{social} = 3.5, SD = 2.1; M_{natural} = 2.7,$

⁵These conditions were also analyzed using Structure (Basic vs. modified) and Negation (yes/no) as fixed effects. Beside the structure effect ($F(1, 2109) = 2945.7, p < 0.001$), there was neither a significant negation effect ($F(1, 2114) = 0.585, p = 0.444$), nor an interaction ($F(1, 2102) = 0.324, p = 0.569$).

SD = 1.8; $F(1, 22) = 5.4$, $p = 0.029$) and without negation ($M_{\text{social}} = 3.5$, SD = 2.1; $M_{\text{natural}} = 2.9$, SD = 1.9; $F(1, 20) = 4.49$, $p = 0.047$).^{6,7}

Focusing on **within-predicate comparisons with adjectives versus nouns**, a mixed model with Entity type (Social vs. Natural) and Predicate type (Adjective vs. Noun) as fixed effects, and participant and item (44 subject + predicate combinations) as random effects, was carried out to test a predicted predicate-type effect and interaction. The test yielded significance for both Predicate type ($F(1, 39) = 164.2$, $p < 0.001$), with higher ratings for comparisons within adjectives ($M = 6.1$, SD = 1.6) than nouns ($M = 4$, SD = 2.1), and Entity type ($F(1, 40) = 8.8$, $p = 0.005$), with higher ratings for comparisons within social items ($M = 5.2$, SD = 2) than within natural items ($M = 4.7$, SD = 2.2). Further, a significant interaction was manifested as predicted ($F(1, 40) = 12.8$, $p = 0.001$). While comparisons with social nouns ($M = 4.5$, SD = 2) were rated higher than with natural nouns ($M = 3.4$, SD = 2), comparisons with social adjectives ($M = 6$, SD = 1.6) and natural adjectives ($M = 6.1$, SD = 1.6) were rated similarly.⁸

Finally, recall that comparative structures were hypothesized to resemble structures with dimensional exception-phrases in involving quantification over dimensions. Thus, the naturalness of a predicate in both types of structure was expected to correlate, as its naturalness in both structures was expected to be affected by the accessibility of its dimensions to binding by quantifiers. Thus, analyses by item were conducted to test correlations. Table 3 presents the **correlation** coefficients and p values for the **comparison by exception-phrase conditions** ($n = 22$). Notice the moderate correlation between the dimensional exception

⁶All **simple structure effects** in comparison constructions were also significant when mixed models were conducted separately for the dataset of natural entity types and for the dataset of social entity types, with structure as the fixed effect.

First, within the natural dataset, structure was significant ($F(2, 775) = 38.5$, $p < 0.001$). The naturalness of the single-subject between-noun comparisons was significantly higher than the naturalness of the within-noun comparisons (Mean difference 0.681, SE = 0.134, $df = 773$, $p < 0.001$), which was significantly higher than the naturalness of the two-subject between-noun comparisons (Mean difference 0.499, SE = 0.127, $df = 776$, $p < 0.001$). Univariate tests based on the pairwise comparisons of marginal means yielded $F(2, 773) = 38.5$, $p < 0.001$.

Second, within the social dataset, structure was significant ($F(2, 826) = 132.9$, $p < 0.001$). The naturalness of the single-subject between-noun comparisons was significantly higher than of the within-noun comparisons (Mean difference 0.665, SE = 0.130, $df = 826$, $p < 0.001$), which was significantly higher than the naturalness of the two-subject between-noun comparisons (Mean difference 1.403, SE = 0.129, $df = 826$, $p < 0.001$). Univariate tests based on the pairwise comparisons of marginal means yielded $F(2, 826) = 132.9$, $p < 0.001$.

By contrast, in the exception phrase modified constructions, structure was significant neither in the natural dataset ($F(1, 434) = 3.28$, $p = 0.071$), nor in the social dataset ($F(1, 528) = 0.22$, $p = 0.64$).

⁷With structure as the fixed effect, the exception phrases (Ex, ExNot) didn't differ from the worst comparison, Be2 s ($F(2, 1554) = 3.4$, $p = 0.034$), but scored lower than the within-noun comparison ($F(2, 1551) = 44.8$, $p < 0.001$).

⁸A mixed model with Than-clause type (phrasal vs. clausal) and predicate type (adjective vs. noun) as fixed effects yielded neither a significant Than-clause type effect ($F(1, 39) = 2.2$, $p = 0.149$), nor an interaction ($F(1, 39) = 3.3$, $p = 0.076$).

Table 3 Correlation coefficients for *exception phrase* by comparison conditions

Spearman's rho		Be1s	ExNot	Within	Be2s
Ex	r^s	0.116	0.701**	0.376	0.241
	p^2	0.6	<0.001	0.084	0.280
Be1s	r^s	1	0.413	0.694**	0.693**
	p^2	.	0.056	<0.001	<0.001
ExNot	r^s		1	0.578**	0.447*
	p^2		.	0.005	0.037
Within	r^s			1	0.452*
	p^2			.	0.035

phrases with negation and both the within-noun comparisons (ExNot \times Within: $r^s = 0.58$, $p = 0.005$) and single-subject between-noun comparison (ExNot \times Be1s: $r^s = 0.45$, $p = 0.037$).

3 Discussion

By the quantificational hypothesis presented in (7), the acceptability of predicates in comparative and exceptive constructions depends on the accessibility of their dimensions. In other words, it depends on the extent to which categorization under the concepts they label can be modeled by means of quantification over their dimensions. Given this hypothesis, a continuum of (un)acceptability was predicted. All nouns were predicted to be less natural than adjectives in within-predicate comparisons, because, by default, their dimensions are bound by averaged similarity operations and thus are not accessible for quantifiers to bind. However, additive nouns were expected to be more acceptable than multiplicative nouns, because under certain circumstances their interpretation turns to one based on dimension counting and quantification. Thus, in particular, additive nouns denoting social categories were expected to be more acceptable than multiplicative nouns denoting animals and plants (Hampton et al. 2009).

The results of this study confirmed the predicted continuum. First, within-predicate comparisons ('{P-er, more P} than y') were rated as more natural with adjectives like *tall* or *big* than with nouns like *a composer* or *pine*.

Second, as Fig. 1 illustrates, additive social nouns were rated as more natural than non-additive natural kind nouns in all the comparison constructions tested. The acceptability of nominal comparisons appears to match the degree to which the noun's default interpretation can be approximated by one based on dimensional quantifiers, as hypothesized. This suggests that these constructions involve operations on accessible dimensions. An accommodation of binary dimensions with equal weights renders the interpretation of additive, but not multiplicative nouns, equivalent to one based on quantification. Thus, the former were predicted to be more acceptable in degree constructions. These predictions were borne out.

Third, similarly, social additive nouns were better than natural-kind multiplicative ones with dimensional exception phrases. The negated condition (as in “not an N except for dim”) indicates the acceptability of dimension binding by an existential quantifier (SOME). By contrast, the non-negated condition (“N except for dim”) indicates the acceptability of dimension binding by a universal quantifier (ALL). The two conditions did not differ significantly, suggesting that neither existential nor universal quantification dominates classification in nouns (as in Shamir 2013). But the more a noun resembles an adjective in its default dimension-binding, the more acceptable it seems to be in either exception phrase construction.

Finally, the correlation matrix in Table 3 suggests that a correlation exists between the naturalness of exception phrase modification and comparison, at least in negated nouns. Comparative structures were hypothesized to resemble structures with dimensional exception-phrases in involving quantification over dimensions. Thus, the naturalness of a predicate in both types of structure was expected to correlate, as its naturalness in both structures was expected to be affected by the accessibility of its dimensions to binding by quantifiers.

The fact that this prediction was only confirmed by the negated condition suggests that compatibility with existential quantification is more important in predicting acceptability of nouns in comparisons. Possibly, comparisons are interpreted by existentially binding the dimensions of their predicative argument. Thus, their naturalness correlates with that of a construction that forces a noun to be interpreted as existential on its dimensions. To assess the validity of this interpretation, future research should assess the naturalness of comparative predicates with dimensional exception phrases (“x is (not) **more P than y** except in dim”). If comparative predicates are predominantly existential, higher acceptability is expected in the negated case.

All in all, the results seem to support the quantifying hypothesis. The study has shown that higher likelihood of categorization based on counting and quantification corresponds with higher naturalness in grammatical constructions involving implicit counting or quantification, and those include not only dimensional exception phrases (e.g., *tiger (in every respect) except for its teeth number*), but also comparisons (e.g., *more a pine than {an oak, that tree, that tree is an oak}*).

However, this study included only positive, one-dimensional, context relative adjectives (e.g., *big*). Future studies should also consider negative (e.g., *small*), absolute (e.g., *empty*, and *closed*), and multidimensional adjectives. The latter are particularly interesting, because they resemble nouns in having multiple dimensions. As such they form a stricter test for the presented theory. They too are expected to be better than nouns in comparative and exceptive structures, since they are expected to have a higher tendency than nouns toward quantificational binding of their dimensions. One multidimensional adjective (*Italian*) appeared in the three comparison types, as part of the 57 fillers. Its averaged scores in the within-predicate comparison *This film is more Italian than that film*, 6.12 (1.21), and two-subject between-predicate comparison *This film is more Italian than that film is American*, 3.92 (1.79), were by far higher than the average scores of the

nouns in these comparisons (4.70 and 2.97), as expected. But its score in the single-subject between-predicate comparison *This film is more Italian than American*, 5.24 (1.68), was similar to the noun scores (4.00). These comparisons appear, indeed, to be especially suited to nouns.

Moreover, this comparison type manifested the least significant correlation. This may stem from the existence of a second reading, as suggested in the discussion of (6d), whereby similarity of an entity to two nominal prototypes is directly compared. This reading does not involve dimension counting, and thus, its contribution to the naturalness judgments is expected to be independent of the naturalness of exception phrases. Its existence is compatible also with the relatively high naturalness of nouns in this construction. At any rate, the clear preference of additive nouns over multiplicative nouns in this comparison does support the existence of a reading based on counting of accessible dimensions, as the paraphrases in (6c–d) suggest.

An anonymous reviewer observed that the between-natural-noun comparisons involved taxonomical hierarchies, e.g., *snake* and *lizard* were predicated over *reptile* and *wolf* and *tiger* over *predator*, whereas some between-social-noun comparisons didn't, e.g., *champion* and *celebrity* were predicated over *football player*. A second reviewer suggested that failure of an implicature may have confounded the results. The suggested implicature is that one of the two compared predicates (e.g., *rose* and *dandelion*) applies to the entity in question, but the speaker does not know which. Naturalness may have been reduced by the difficulty of imagining a speaker being unsure of whether a flower is a rose or a dandelion. These potential confounds deserve investigation (e.g., by presenting sentences in contexts where the implicature is either satisfied or not), but they cannot explain why differences between natural and social nouns occurred also in several other constructions, specifically within-noun comparisons and exception phrases, of which the between-noun comparisons were actually more acceptable.

A number of additional issues merit attention. Recall that the embedded clause in the within-predicate comparisons was as in *This football player is more a champion than that football player (is)*, rather than *than that one (is)*. The idea was to block interpretations of *that one* as relating to a property, which would turn the comparison into a between-noun one, but this feature may have reduced the naturalness of these comparisons, and thus should be eliminated in future studies. Moreover, in the exception phrase conditions, each noun was matched with but one dimension. This fact may have added noise that reduced the significance of the results. Future research should assess judgments for exception phrases with at least 3–4 dimensions per noun or adjective (cf., Sassoon 2012. and Shamir 2013), as well as with other quantificational constructions, such as *P in every respect, some respect, and most respects*. Nouns and adjectives of additional domains can be tested (for instance, abstract nouns: *problem, love*), as well as additive natural kind nouns and multiplicative social nouns, assuming such exist. The role of mediating particles (cf., *of* in (3a)) should also be addressed.

The present study has broader implications for our understanding of the adjective-noun word class distinction, including in particular 'nouny' adjectives and adjective-like nouns. Consider in particular, nationality concepts such as *American*

or trait concepts such as *extrovert*. They can fill in attributive-adjective positions as in *extrovert personality* and *American nationality*, but unlike most adjectives, they can also ‘stand alone’ in argument positions without any explicit noun around, as in *Americans* and *extroverts*, and unlike the ill-formed *#talls* and *#healthies*. An explanation in terms of categorization based on accessible dimensions deserves consideration. Possibly, in argument position, the dimensions are bound by additive similarity functions, with equal dimensional weights. But since this type of interpretation can be represented by means of a dimension set bound by a quantifier, these nouns easily occupy adjectival positions.

Additional implications pertain to size adjectives such as *big* and *huge*. When modifying certain nouns, as in *this midget is a big fan*, they do not attribute big physical size. Morzycki (2009) argues that this interpretation of *big* occurs when it modifies a noun that is itself morphologically gradable (e.g., *fan*, *stamp-collector*, *genius* or *idiot*). For example, *smoker* denotes a measurement of frequency or affinity for smoking. On this view, size adjectives function as degree modifiers, similar to *very* (type < d,t >); e.g., *big* denotes a mapping f_{big} of nominal degrees *d* to their relative size, $f_{\text{big}}(d)$. Morzycki further observes that nominal gradability readings of size adjectives are absent in predicate position (as in *the fan is big*), or in negative size adjectives (as in *a small fan just came in*, which relates to physical size, not fan-hood). These features are unique to nominal-gradability readings, as opposed to abstract size readings (as in *this mistake is big/small*) and significance readings (*he is huge!*). The position generalization is captured by postulating a morpheme *meas* in the nominal degree projection, which is not available in other positions. *Meas* takes as arguments a gradable noun, a size adjective, and an entity and returns truth if and only if the entity falls under the noun and its degree in the noun is big.

However, this account of gradable nouns in terms of degree functions fails to explain why these nouns are not generally compatible with gradability morphemes (e.g., **faner*, **fanest*, **too fan*), and why their combinations with *more* must be mediated, as in *more of an idiot than Bill*. The results reported in this paper give rise to an alternative approach, which rests on the observation that the gradable dimensions of adjectives are often context dependent, and their choice is affected, among other things, by the noun they modify. For example, the dimension of *long* is different for tables than for stories. Assuming that nouns are themselves associated with dimension sets, size adjectives may, under certain circumstances, be able to access and operate on these sets. The position generalization can be captured by postulating that, unless an adjective has a dimension set argument (as *typical* does in, e.g., *typical of birds*), only in attributive position can it access the dimension set of the noun it modifies (perhaps by virtue of a syntactic head whose semantics does the job). This account uses Morzycki’s logic, but spares the need to stipulate that some nouns denote a degree function, while most other nouns do not, or that size adjectives are adverbial (adnominal) modifiers, while most other adjectives are not.

This approach extends beyond size adjectives and explains the different interpretations of *old* in *He is an old friend* and *This friend is old*. Thus, the availability of nominal dimensions for adjectives to operate on is a more general phenomenon in attributive position. Moreover, the set of examples of gradable nouns reported in

the literature typically includes nouns such as *idiot*, *nerd*, *soccer fan*, *airhead*, *goat cheese enthusiast*, *simpleton*, *Barbie doll lover*, *loser*, and *weirdo* (de Vries 2010). It is easy to observe that all of them belong to the social domain. Thus, their dimensions are expected to be relatively accessible for grammatical operations to quantify over as in *idiot with respect to his political views* and *complete/total idiot* (de Vries 2010). Moreover, these nouns seem to have morphologically gradable dimensions (e.g., *stupid*, *intelligent*, *admiring*) accessible for size adjectives to select and use for their interpretation.

Dimension accessibility for quantifier binding seems to affect additional gradable constructions. For example, in various languages human categories such as *boy* and *girl* directly combine with the modifier *very*, as in the Hebrew *meod yalada* ('very girl', 'very girlish/immature') and Spanish *Es muy hombre* ('is very man', 'He is very much a man'; Espinal 2013). Additive dimension binding for the noun *girl* makes its dimensions more accessible than those of most other nouns. The importance of gradable traits in the stereotype of girls (e.g., loving pink) can be stretched to the point that *girl* is interpreted as equivalent to *girlish*, thereby licensing *very*. In sum, nouns carrying expressive or evaluative components such as *idiot*, *coward*, *hero* or *child* in its metaphoric sense, get as close to gradable adjectives as nouns can (Constantinescu 2011: 49–96), perhaps because they have adjectival dimensions (another point to consider in the future).

To conclude, the main result of the reported study is a connection between the acceptability of a given noun or adjective in comparison constructions and its type of characteristic categorization criterion (i.e., whether, as a default, its dimensions combine into a single criterion via quantifiers or other operations). This result highlights connections between cognitive psychological findings and linguistic phenomena, thus potentially contributing to the study of morphological gradability within linguistics and to an improved understanding of certain experimental results in cognitive psychology, where the potential role of the noun-adjective distinction was overlooked (Wattenmaker 1995). By raising awareness both to grammatical and conceptual distinctions, and by pointing out directions for future research, this paper has aimed to deepen our understanding of the relations between the formal and conceptual components of natural languages.

Acknowledgements Many thanks for the financial support by a Grant from the GIF, the German-Israeli Foundation for Scientific Research and Development. Special thanks also to the editors, James Hampton and Yoadd Winter, for the inspiring workshop and very useful comments.

Appendix A: Averages on 25 Participants for the Comparison Conditions with Nouns (Left) and Adjectives (Right)

Nominal predicates by entity type	Single subject between-noun comparisons		Within-noun comparisons		Two-subject between-noun comparisons		Adjectival predicates by entity type	Within-adjective comparisons	
	More NP1 than NP2		More NP than Y (is)		More NP ¹ than y is NP ²			More AP than Y (is)	
Animals	M	SD	M	SD	M	SD	Animals	M	SD
Cow	3.64	2.06	2.92	2.00	2.32	1.62	Taller	5.88	1.56
Mosquito	2.64	1.94	2.68	1.76	2.52	1.70	Bigger	6.84	0.61
Crow	2.68	2.03	2.16	1.54	1.96	1.37	More expensive	6.32	1.12
Snake	4.84	1.93	3.6	1.70	3.16	2.13	More colorful	6.72	0.60
Tiger	4.28	2.03	3.28	1.80	2.84	1.69	Heavier	5.6	1.74
Dolphin	3.6	2.15	2.8	1.74	2.56	1.24	Older	5.48	2.14
Plants							Plants		
Pine	4.52	1.96	4.32	1.93	3.28	1.93	Older	5.84	1.83
Potato	3.16	2.19	2.12	1.45	2.2	1.50	Bigger	6.4	0.94
Dandelion	3.72	1.97	4.08	2.00	2.8	1.60	More expensive	6.08	1.65
Mint	3.32	1.89	3.76	1.94	2.52	1.65	More colorful	6.72	0.72
Moss	4.76	1.97	3.52	1.90	3.4	1.79	Taller	5.68	1.76
Orange	3.6	2.00	3.4	2.23	2.48	1.58	Heavier	6.4	1.39
Artifacts							Artifacts		
Church	5.48	1.60	4.48	1.90	4.24	2.06	Older	6.16	1.59
Dress	5.44	1.63	3.8	2.15	2.8	1.83	More colorful	6	1.52
Taxi	3.72	2.20	4	1.94	2.92	1.65	Heavier	5.88	1.70
Theatre	5.04	2.05	4.08	2.26	3.44	1.83	Bigger	6.6	0.69
Vase	5.76	1.75	4.32	1.69	3.16	1.87	Taller	6.24	1.14
Diary	6.2	1.26	4.44	1.92	3.32	1.93	More expensive	6.64	1.02
Humans							Humans		
Reporter	5.48	1.50	4.12	1.99	2.96	1.71	Older	5.4	1.85
Composer	5.76	1.36	4.76	2.03	3.12	1.68	More colorful	5.68	1.64
Champion	5.44	1.36	4.76	1.77	3.56	1.83	More expensive	4.68	2.26
Genius	4.32	2.22	5.24	1.75	2.12	1.53	Bigger	6.6	0.94
Businesswoman	4.88	1.92	4.32	1.87	3.16	1.97	Heavier	5.56	1.55
Person	5.2	2.14	5.2	2.04	3	2.02	Taller	6.8	0.63
M	4.48	1.03	3.84	0.86	2.91	0.52	M	6.09	0.54
SD	1.03	0.28	0.86	0.20	0.52	0.22	SD	0.54	0.50

Appendix B: Averages on 25 Participants for the 24 Nouns in Baseline and Modified Conditions

Entity type	Item Noun	Basic categorization conditions (Pos)		Negated categorization conditions (Not)		Exception phrase condition (Ex)		Negated exception phrase condition (ExNot)	
		M	SD	M	SD	M	SD	M	SD
Animals	Cow	6.68	1.12	6.68	0.68	1.92	1.29	2.04	1.40
	Mosquito	3.48	2.44	5	2.43	2.24	1.70	2.88	1.90
	Crow	3.04	2.51	3.44	2.56	1.8	1.30	1.68	1.41
	Snake	6.72	0.87	6.64	1.23	3.64	2.31	4	1.94
	Tiger	6.52	1.06	6.76	0.65	2.32	1.49	1.96	1.31
	Dolphin	6.68	0.68	6.36	1.23	3.08	1.79	3.52	1.96
Plants	Pine	6.64	1.09	6.76	0.59	2.84	1.38	2.64	1.57
	Potato	6.64	0.93	6.32	1.64	2.88	1.99	2	1.50
	Dandelion	6.84	0.46	6.92	0.39	3.28	2.16	2.68	1.67
	Mint	5.12	1.88	5.12	2.16	3.12	2.10	2.52	1.53
	Moss	5.92	1.92	6.72	0.72	2.48	1.86	2.76	2.12
	Orange	6.68	0.88	6.64	0.79	3.04	1.93	2.48	1.68
Artifacts	Church	6.92	0.27	6.72	0.66	4.32	2.48	5.24	2.34
	Dress	6.72	0.60	6.84	0.54	2.88	1.88	2.84	2.07
	Taxi	6.52	1.02	6.84	0.54	3.24	1.94	3.48	1.94
	Theatre	6.12	1.77	6.12	1.68	4.32	2.11	3.64	2.02
	Vase	6.72	0.66	6.52	1.30	3.04	1.78	2.16	1.57
	Diary	6.16	1.01	6.92	0.27	4.08	1.74	3.56	2.14
Humans	Reporter	6.6	1.10	6	1.62	2.56	1.88	2.4	1.50
	Composer	6.8	0.63	6.84	0.61	2.72	1.89	3.84	2.22
	Champion	7	0.00	6.76	0.59	4.28	1.76	4.28	1.84
	Genius	6.84	0.61	6.88	0.43	5.04	1.99	5.16	1.57
	Businesswoman	5.72	1.25	5.16	1.87	3.08	1.57	2.76	1.53
	Person	6.8	0.80	6.84	0.46	2.56	1.90	3.08	1.81
	M	6.25	1.07	6.32	1.07	3.11	0.81	3.07	0.95
	SD	1.02	0.63	0.84	0.69	0.81	0.29	0.95	0.29

Appendix C: The Full Set of Sentences Together with Their Naturalness Mean and Sd

Item	Structure: basic form	Mean	Std
1.	[BaA1] This farm animal is a cow	6.68	1.12
2.	[BaA2] This fly is a mosquito	3.48	2.44
3.	[BaA3] This bird is a crew	3.04	2.51
4.	[BaA4] This reptile is a snake	6.72	0.87
5.	[BaA5] This predator is a tiger	6.52	1.06
6.	[BaA6] This aquatic mammal is a dolphin	6.68	0.68
7.	[BaH1] This journalist is a reporter	6.6	1.10
8.	[BaH2] This artist is a composer	6.8	0.63
9.	[BaH3] This football player is a champion	7	0.00
10.	[BaH4] This girl is a genius	6.84	0.61
11.	[BaH5] This colleague is a businesswoman	5.72	1.25
12.	[BaH6] This person is a victim	6.8	0.80
13.	[BaP1] This tree is a pine	6.64	1.09
14.	[BaP2] This vegetable is a potato	6.64	0.93
15.	[BaP3] This flower is a dandelion	6.84	0.46
16.	[BaP4] This spice is a mint	5.12	1.88
17.	[BaP5] This plant is moss	5.92	1.92
18.	[BaP6] This piece of fruit is an orange	6.68	0.88
19.	[BaT1] This place is a church	6.92	0.27
20.	[BaT2] This piece of clothing is a dress	6.72	0.60
21.	[BaT3] This car is a taxi	6.52	1.02
22.	[BaT4] This hall is a theatre	6.12	1.77
23.	[BaT5] This container is a vase	6.72	0.66
24.	[BaT6] This booklet is a diary	6.16	1.01
	Structure: Negated Basic form	6.25 (1.02)	1.07 (0.63)
1.	[BaNA1] This farm animal is not a cow	6.68	0.68
2.	[BaNA2] This fly is not a mosquito	5	2.43
3.	[BaNA3] This bird is not a crew	3.44	2.56
4.	[BaNA4] This reptile is not a snake	6.64	1.23
5.	[BaNA5] This predator is not a tiger	6.76	0.65
6.	[BaNA6] This aquatic mammal is not a dolphin	6.36	1.23
7.	[BaNH1] This journalist is not a reporter	6	1.62
8.	[BaNH2] This artist is not a composer	6.84	0.61
9.	[BaNH3] This football player is not a champion	6.76	0.59
10.	[BaNH4] This girl is not a genius	6.88	0.43
11.	[BaNH5] This colleague is not a businesswoman	5.16	1.87
12.	[BaNH6] This person is not a victim	6.84	0.46
13.	[BaNP1] This tree is not a pine	6.76	0.59
14.	[BaNP2] This vegetable is not a potato	6.32	1.64

(continued)

(continued)

Item	Structure: basic form	Mean	Std
15.	[BaNP3] This flower is not a dandelion	6.92	0.39
16.	[BaNP4] This spice is not a mint	5.12	2.16
17.	[BaNP5] This plant is not moss	6.72	0.72
18.	[BaNP6] This piece of fruit is not an orange	6.64	0.79
19.	[BaNT1] This place is not a church	6.72	0.66
20.	[BaNT2] This piece of clothing is not a dress	6.84	0.54
21.	[BaNT3] This car is not a taxi	6.84	0.54
22.	[BaNT4] This hall is not a theatre	6.12	1.68
23.	[BaNT5] This container is not a vase	6.52	1.30
24.	[BaNT6] This booklet is not a diary	6.92	0.27
	Structure: Single entity Between noun comparisons	6.32 (0.84)	1.07 (0.07)
1.	[Be1s1A1] This farm animal is more a cow than a horse	3.64	2.06
2.	[Be1s1A2] This fly is more a mosquito than a wasp	2.64	1.94
3.	[Be1s1A3] This bird is more a crow than a pigeon	2.68	2.03
4.	[Be1s1A4] This reptile is more a snake than a lizard	4.84	1.93
5.	[Be1s1A5] This predator is more a wolf than a tiger	4.28	2.03
6.	[Be1s1A6] This aquatic mammal is more a dolphin than a shark	3.6	2.15
7.	[Be1sP1] This tree is more a pine than an oak	4.52	1.96
8.	[Be1sP2] This vegetable is more a potato than a carrot	3.16	2.19
9.	[Be1sP3] This flower is more a dandelion than a rose	3.72	1.97
10.	[Be1sP4] This spice is more onion than mint	3.32	1.89
11.	[Be1sP5] This plant is more moss than grass	4.76	1.97
12.	[Be1sP6] This piece of fruit is more an orange than an apple	3.6	2.00
13.	[Be1sT1] This place is more a church than an art gallery	5.48	1.60
14.	[Be1sT2] This piece of clothing is more a dress than a nightshirt	5.44	1.63
15.	[Be1sT3] This car is more a taxi than an ambulance	3.72	2.20
16.	[Be1sT4] This hall is more a theatre than a cinema	5.04	2.05
17.	[Be1sT5] This container is more a carafe than a vase	5.76	1.75
18.	[Be1sT6] This booklet is more a diary than a sketchbook	6.2	1.26
19.	[Be1sH1] This journalist is more a consultant than a reporter	5.48	1.50
20.	[Be1sH2] This artist is more a composer than a poet	5.76	1.36
21.	[Be1sH3] This football player is more a champion than a celebrity	5.44	1.36
22.	[Be1sH4] This girl is more a genius than a child	4.32	2.22
23.	[Be1sH5] This colleague is more a businesswoman than a friend	4.88	1.92
24.	[Be1sH6] This person is more a victim than a witness	5.2	2.14
	Structure: Two entity Between noun comparisons	4.48 (1.03)	1.03 (0.28)
1.	[Be2sA1] This farm animal is more a cow than that farm animal is a horse	2.32	1.62
2.	[Be2sA2] This fly is more a mosquito than that fly is a wasp	2.52	1.70
3.	[Be2sA3] This bird is more a crow than that bird is a pigeon	1.96	1.37
4.	[Be2sA4] This reptile is more a snake than that reptile is a lizard	3.16	2.13
5.	[Be2sA5] This predator is more a wolf than that predator is a tiger	2.84	1.69

(continued)

(continued)

Item	Structure: basic form	Mean	Std
6.	[Be2sA6] This aquatic mammal is more a dolphin than that aquatic mammal is a shark	2.56	1.24
7.	[Be2sP1] This tree is more a pine than that tree is an oak	3.28	1.93
8.	[Be2sP2] This vegetable is more a potato than that vegetable a carrot	2.2	1.50
9.	[Be2sP3] This flower is more a dandelion than that flower is a rose	2.8	1.60
10.	[Be2sP4] This spice is more onion than that spice is a mint	2.52	1.65
11.	[Be2sP5] This plant is more moss than that plant is grass	3.4	1.79
12.	[Be2sP6] This piece of fruit is more an orange than that piece of fruit is an apple	2.48	1.58
13.	[Be2sT1] This place is more a church than that place is an art gallery	4.24	2.06
14.	[Be2sT2] This piece of clothing is more a dress than that piece of clothing is a nightshirt	2.8	1.83
15.	[Be2sT3] This car is more a taxi than that car is an ambulance	2.92	1.65
16.	[Be2sT4] This hall is more a theatre than that hall is a cinema	3.44	1.83
17.	[Be2sT5] This container is more a carafe than that container is a vase	3.16	1.87
18.	[Be2sT6] This booklet is more a diary than that booklet is a sketchbook	3.32	1.93
19.	[Be2sH1] This journalist is more a consultant than that journalist is a reporter	2.96	1.71
20.	[Be2sH2] This artist is more a composer than that artists is a poet	3.12	1.68
21.	[Be2sH3] This football player is more a champion than that football player is a celebrity	3.56	1.83
22.	[Be2sH4] This girl is more a genius than that girl is a child	2.12	1.53
23.	[Be2sH5] This colleague is more a businesswoman than that colleague is a friend	3.16	1.97
24.	[Be2sH6] This person is more a victim than that person is a witness	3	2.02
	Structure: Within adjective comparison	2.91 (0.52)	0.52 (0.22)
25.	[WaA1] This farm animal is taller than that farm animal is	5.88	1.56
26.	[WaA2] This fly is bigger than that fly	6.84	0.61
27.	[WaA3] This bird is more expensive than that bird is	6.32	1.12
28.	[WaA4] This reptile is more colorful than that reptile	6.72	0.60
29.	[WaA5] This predator is heavier than that predator is	5.6	1.74
30.	[WaA6] This aquatic mammal is older than that aquatic mammal	5.48	2.14
31.	[WaP1] This tree is older than that tree is	5.84	1.83
32.	[WaP2] This vegetable is bigger than that vegetable	6.4	0.94
33.	[WaP3] This flower is more expensive than that flower is	6.08	1.65
34.	[WaP4] This spice is more colorful than that spice	6.72	0.72
35.	[WaP5] This plant is taller than that plant is	5.68	1.76
36.	[WaP6] This piece of fruit is heavier than that piece of fruit	6.4	1.39
37.	[WaT1] This place is older than that place is	6.16	1.59

(continued)

(continued)

Item	Structure: basic form	Mean	Std
38.	[WaT2] This piece of clothing is more colorful than that piece of clothing	6	1.52
39.	[WaT3] This car is heavier than that car is	5.88	1.70
40.	[WaT4] This hall is bigger than that hall	6.6	0.69
41.	[WaT5] This container is taller than that container is	6.24	1.14
42.	[WaT6] This booklet is more expensive than that booklet	6.64	1.02
43.	[WaH1] This journalist is older than that journalist is	5.4	1.85
44.	[WaH2] This artist is more colorful than that artist	5.68	1.64
45.	[WaH3] This football player is more expensive than that football player is	4.68	2.26
46.	[WaH4] This girl is bigger than that girl	6.6	0.94
47.	[WaH5] This colleague is heavier than that colleague is	5.56	1.55
48.	[WaH6] This person is taller than that person	6.8	0.63
	Structure: Within noun comparisons	6.09 (0.54)	0.54 (0.50)
1.	[WhA1] This farm animal is more a cow than that farm animal is	2.92	2.00
2.	[WhA2] This fly is more a mosquito than that fly	2.68	1.76
3.	[WhA3] This bird is more a crow than that bird is	2.16	1.54
4.	[WhA4] This reptile is more a snake than that reptile	3.6	1.70
5.	[WhA5] This predator is more a tiger than that predator is	3.28	1.80
6.	[WhA6] This aquatic mammal is more a dolphin than that aquatic mammal	2.8	1.74
7.	[WhP1] This tree is more a pine than that tree is	4.32	1.93
8.	[WhP2] This vegetable is more a potato than that vegetable	2.12	1.45
9.	[WhP3] This flower is more a dandelion than that flower is	4.08	2.00
10.	[WhP4] This spice is more a mint than that spice	3.76	1.94
11.	[WhP5] This plant is more moss than that plant is	3.52	1.90
12.	[WhP6] This piece of fruit is more an orange than that piece of fruit	3.4	2.23
13.	[WhT1] This place is more a church than that place is	4.48	1.90
14.	[WhT2] This piece of clothing is more a dress than that piece of clothing	3.8	2.15
15.	[WhT3] This car is more a taxi than that car is	4	1.94
16.	[WhT4] This hall is more a theatre than that hall	4.08	2.26
17.	[WhT5] This container is more a vase than that container is	4.32	1.69
18.	[WhT6] This booklet is more a diary than that booklet	4.44	1.92
19.	[WhH1] This journalist is more a reporter than that journalist is	4.12	1.99
20.	[WhH2] This artist is more a composer than that artist	4.76	2.03
21.	[WhH3] This football player is more a champion than that football player is	4.76	1.77
22.	[WhH4] This girl is more a genius than that girl	5.24	1.75
23.	[WhH5] This colleague is more a businesswoman than that colleague is	4.32	1.87
24.	[WhH6] This person is more a victim than that person	5.2	2.04

(continued)

(continued)

Item	Structure: basic form	Mean	Std
	Structure: Modified forms	3.84 (0.86)	0.86 (0.20)
1.	[ExA1] This farm animal is a cow except for its milk taste	1.92	1.29
2.	[ExA2] This fly is a mosquito except with respect to the chemistry of its blood	2.24	1.70
3.	[ExA3] This bird is a crew except for its eating habits	1.8	1.30
4.	[ExA4] This reptile is a snake except with respect to eye structure	3.64	2.31
5.	[ExA5] This predator is a tiger except for its teeth number	2.32	1.49
6.	[ExA6] This aquatic mammal is a dolphin except in behavior	3.08	1.79
7.	[Exp1] This tree is a pine except with respect to the wood fibers	2.84	1.38
8.	[Exp2] This vegetable is a potato except for color	2.88	1.99
9.	[Exp3] This flower is a dandelion except in size	3.28	2.16
10.	[Exp4] This spice is a mint except for the bitterness	3.12	2.10
11.	[Exp5] This plant is moss except with respect to metabolism	2.48	1.86
12.	[Exp6] This piece of fruit is an orange except for shape	3.04	1.93
13.	[Ext1] This place is a church except in appearance	4.32	2.48
14.	[Ext2] This piece of clothing is a dress except for its use	2.88	1.88
15.	[Ext3] This car is a taxi except for its current function	3.24	1.94
16.	[Ext4] This hall is a theatre except with respect to seat type	4.32	2.11
17.	[Ext5] This container is a vase except with respect to texture	3.04	1.78
18.	[Ext6] This booklet is a diary except for the sketches	4.08	1.74
19.	[ExH1] This journalist is a reporter except for salary	2.56	1.88
20.	[ExH2] This artist is a composer except with respect to education	2.72	1.89
21.	[ExH3] This football player is a champion except for running abilities	4.28	1.76
22.	[ExH4] This girl is a genius except with respect to literature.	5.04	1.99
23.	[ExH5] This colleague is a businesswoman except in generosity.	3.08	1.57
24.	[ExH6] This person is a victim except for subjective experience	2.56	1.90
	Structure: Negated modified forms	3.11 (0.81)	0.81 (0.29)
1.	[ExNA1] This farm animal is not a cow except for its milk taste	2.04	1.40
2.	[ExNA2] This fly is not a mosquito except with respect to the chemistry of its blood	2.88	1.90
3.	[ExNA3] This bird is not a crew except for its eating habits	1.68	1.41
4.	[ExNA4] This reptile is not a snake except with respect to eye structure	4	1.94
5.	[ExNA5] This predator is not a tiger except for its teeth number	1.96	1.31
6.	[ExNA6] This aquatic mammal is not a dolphin except in behavior	3.52	1.96
7.	[ExNP1] This tree is not a pine except with respect to the wood fibers	2.64	1.57
8.	[ExNP2] This vegetable is not a potato except for color	2	1.50
9.	[ExNP3] This flower is not a dandelion except in size	2.68	1.67
10.	[ExNP4] This spice is not a mint except for the bitterness	2.52	1.53

(continued)

(continued)

Item	Structure: basic form	Mean	Std
11.	[ExNP5] This plant is not moss except with respect to metabolism	2.76	2.12
12.	[ExNP6] This piece of fruit is not an orange except for shape	2.48	1.68
13.	[ExNT1] This place is not a church except in appearance	5.24	2.34
14.	[ExNT2] This piece of clothing is not a dress except for its use	2.84	2.07
15.	[ExNT3] This car is not a taxi except for its current function	3.48	1.94
16.	[ExNT4] This hall is not a theatre except with respect to seat type	3.64	2.02
17.	[ExNT5] This container is not a vase except with respect to texture	2.16	1.57
18.	[ExNT6] This booklet is not a diary except for the sketches	3.56	2.14
19.	[ExNH1] This journalist is not a reporter except for salary	2.4	1.50
20.	[ExNH2] This artist is not a composer except with respect to education	3.84	2.22
21.	[ExNH3] This football player is not a champion except for running abilities	4.28	1.84
22.	[ExNH4] This girl is not a genius except with respect to literature.	5.16	1.57
23.	[ExNH5] This colleague is not a businesswoman except in generosity.	2.76	1.53
24.	[ExNH6] This person is not a victim except for subjective experience	3.08	1.81
	Fillers	3.07 (0.95)	0.95 (0.29)
1.	[Filba] This chair is often wide	2.28	2.01
2.	[Filba] This inspector saw anything	1.96	1.46
3.	[Filgo] This inspector did not see anything	6.8	0.63
4.	[Filgo] This chair is often covered	6.16	1.29
5.	[HeBa] This boy is healthy	6.88	0.43
6.	[HeEx] This boy is healthy except for a cold	4.88	1.90
7.	[HeExN] This boy is not healthy except for a cold	1.52	1.17
8.	[HeNBa] This boy is sick	6.92	0.27
9.	[HeNEx] This boy is sick except for a cold	1.48	0.98
10.	[HeNExN] This boy is not sick except for a cold	4.16	2.07
11.	[NaBa] This film is Italian	6.96	0.20
12.	[NaBa] This immigrant is Italian	6.92	0.27
13.	[NaBe1s] This film is more Italian than American	6.12	1.21
14.	[NaBe2s] This film is more Italian than that film is American	3.92	1.79
15.	[NaCo1] This film is Italian compared to that film	3.2	1.96
16.	[NaCo2] This man is Italian compared to that film	1.92	1.35
17.	[NaEx1] This film is Italian except with respect to the music	4.84	1.99
18.	[NaEx2] This immigrant is an Italian except with respect to the accent	3.08	2.00
19.	[NaExN1] This film is not Italian except with respect to the music	4.52	2.04
20.	[NaExN2] This immigrant is not an Italian except with respect to the accent	3.48	2.23
21.	[NaWh1] This film is more Italian than that film	5.24	1.68
22.	[Filba] This musician is a big fan	5.24	1.86

(continued)

(continued)

Item	Structure: basic form	Mean	Std
23.	[Filba] This musician is a small fan	2.84	1.78
24.	[Filba] This musician is a bigger fan than that musician	4.2	2.14
25.	[Filba] This musician is a smaller fan than that musician	2.44	1.70
26.	[Cob1] This banana is bent compared to that banana	5.12	2.20
27.	[Cob2] This rod is bent compared to that rod	6.16	1.49
28.	[Cob3] This bed is bent compared to that bed	5.08	2.02
29.	[Coba1] This creature is bald compared to that creature	5.76	1.03
30.	[Cobi1] This building is big compared to that building	6.56	0.90
31.	[Cobi2] This pencil is big compared to that pencil	6.68	0.61
32.	[Coch1] This dress is cheap compared to that dress	6.84	0.46
33.	[Coch2] This apartment is cheap compared to that apartment	6.92	0.27
34.	[Coe1] This tea glass is empty compared to that tea glass	5.88	1.37
35.	[Coe2] This wine glass is empty compared to that wine glass	5.64	1.62
36.	[Coe3] This espresso cup is empty compared to that espresso cup	4.56	1.94
37.	[Coe4] This jam jar is empty compared to that jam jar	5.36	1.81
38.	[Coe5] This whiskey glass is empty compared to that whiskey glass	5.36	1.85
39.	[Coex1] This dress is expensive compared to that dress	6.88	0.32
40.	[Coex2] This apartment is expensive compared to that apartment	6.8	0.40
41.	[Cof1] This tea glass is full compared to that tea glass	5.6	1.67
42.	[Cof2] This wine glass is full compared to that wine glass	5.92	1.38
43.	[Cof3] This espresso cup is full compared to that espresso cup	5.64	1.65
44.	[Cof4] This jam jar is full compared to that jam jar	5.72	1.37
45.	[Cof5] This whiskey glass is full compared to that whiskey glass	6	1.26
46.	[Coh1] This creature is hairy compared to that creature	6.48	0.81
47.	[Cos1] This banana is straight compared to that banana	5.92	1.57
48.	[Cos2] This rod is straight compared to that rod	6.04	1.56
49.	[Cos3] This bed is straight compared to that bed	5.92	1.32
50.	[Cosh1] This girl is short compared to that girl	6.56	0.80
51.	[Cosh2] This basketball player is short compared to that basketball player	6.56	0.94
52.	[Cosm1] This building is small compared to that building	6.8	0.49
53.	[Cosm2] This pencil is small compared to that pencil	6.52	0.90
54.	[Cot1] This girl is tall compared to that girl	6.28	1.28
55.	[Cot2] This basketball player is tall compared to that basketball player	6.28	0.87
56.	[Cota1] This ice-cream is tasty compared to that ice-cream	6.56	0.64
57.	[Cotle1] This ice-cream is tasteless compared to that ice-cream	6.6	0.89

References

- Baker, M. C. (2003). *Lexical categories: Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.
- Bard, E. G., Robertson, D., & Sorace A. (1996). Magnitude estimation of linguistic acceptability. *Language*, 72(1), 32–68.
- Bartsch, R. (1986). Context-dependent interpretations of lexical items. In J. Groenendijk, D. de Jongh & M. Stokhof (Eds.), *Foundations of pragmatics and lexical semantics*, GRASS 7 (pp. 1–26). Dordrecht: Foris.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2006). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3–5.
- Constantinescu, C. (2011). Gradability in the nominal domain. Doctoral dissertation, Rijksuniversiteit Leiden.
- Davies, M. (2010). Corpus of Contemporary American English (<http://corpus.byu.edu/coca>). Brigham Young University.
- de Vries, H. (2010). Evaluative degree modification of adjectives and nouns. Master's thesis, Utrecht Institute of Linguistics.
- Espinal, M. T. (2013). Bare nominals, bare predicates: Properties and related types. In J. Kabatek & A. Wall (Eds.), *Bare noun phrases in romance: Theory and (empirical) data*. Amsterdam: John Benjamins.
- Fox, D., & Hackl, M. (2006). The universal density of measurement. *Linguistics and Philosophy*, 29, 537–586.
- Gardenfors, P. (2004). *Conceptual spaces—The geometry of thought*. Cambridge, Massachusetts: MIT Press.
- Giannakidou, A., & Yoon, S. (2011). The subjective mode of comparison: Metalinguistic comparatives in Greek and Korean. *Natural Language & Linguistic Theory*, 29, 621–655.
- Hampton, J. A., Storms, G., Simmons, C. L., & Heussen, D. (2009). Feature integration in natural language concepts. *Memory and Cognition*, 37(8), 1150–1163.
- Hampton, J. A. (2007). Typicality, graded membership and vagueness. *Cognitive Science*, 31, 355–383.
- Hampton, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory and Language*, 34, 686–708.
- Hampton, J. A. (1998). Similarity based categorization and fuzziness of natural categories. *Cognition*, 65, 137–165.
- Heim, I. (2000). Degree operators and scope. In *Proceedings of SALT 10*. Ithaca, NY: CLC Publications.
- Hoeksema, J. (1995). The semantics of exception phrases. In J. van Eijck & J. van der Does (Eds.), *Quantifiers, logic, and language* (pp. 145–177). Stanford: CSLI.
- Kamp, H. (1975). Two theories about adjectives. In E. L. Keenan (Ed.), *Formal semantics of natural language* (pp. 123–155). Cambridge: Cambridge University Press.
- Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland.
- Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4(1), 1–45.
- Landman, F. (1989). Groups II. *Linguistics and Philosophy*, 12, 723–744.
- Lasnik, P. (1999). Pragmatic halos. *Language*, 75, 522–551.
- McCloskey, M. E., & Glucksberg, S. (1978). 'Natural categories: Well defined or fuzzy sets'? *Memory & Cognition*, 6, 462–472.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85(3), 207–238.

- Moltmann, F. (1995). Exception phrases and polyadic quantification. *Linguistics and Philosophy*, 18, 223–280.
- Morzycki, M. (2009). Degree modification of gradable nouns: Size adjectives and adnominal degree morphemes. *Natural Language Semantics*, 17(2), 175–203.
- Morzycki, M. (2011). Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics*, 19(1), 39–86.
- Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: The MIT Press.
- Rosch, E. (1973). Natural categories. *Cognitive Psychology*, 4(3), 328–350.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Roth, E., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology*, 15, 346–378.
- Sassoon, G. W. (2012). The double nature of negative antonymy. In *The Proceedings of Sinn und Bedeutung* 16. Utrecht.
- Sassoon, G. W. (2013a). *Vagueness, gradability, and typicality, the interpretation of adjectives and nouns*. Leiden: Brill.
- Sassoon, G. W. (2013b). A typology of multidimensional adjectives. *Journal of Semantics*, 30, 335–380.
- Sassoon, G. W. (2015). Between noun comparisons. In *Selected papers from the 10th TbiLLC conference, Lecture notes in artificial intelligence*. Springer.
- Sassoon, G. W. (2014). The type of adjectives. In M. Aloni, M. Franke & F. Roelofsen (Eds.), *Proceedings of the 19th Amsterdam colloquium*.
- Shamir, A. (2013) Disjunctivity. Master thesis. The Hebrew University, Jerusalem.
- Tversky, A. (1977). Dimensions of similarity. *Psychological Review*, 84, 327–352.
- von Fintel, K. (1994). Restrictions on quantifier domains. Doctoral dissertation, University of Massachusetts, Amherst.
- von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of Semantics*, 3, 1–77.
- Wattenmaker, W. D. (1995). Knowledge structures and linear separability: Integrating information in object and social categorization. *Cognitive Psychology*, 28, 274–328.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Author Index

A

Agid, Yves, 216
Aina, B., 110
Allen, S. W., 16
Alonso, Jose-Manuel, 217
Anderson, A., 20
Andersson, J. M., 110
Andrew, J. M., 100
Anishchanka, A., 126
Aquinas, 224, 234, 236, 238, 240
Arbuckle, Robert, 216
Aristotle, 224, 234, 236, 238, 240
Armstrong, S. L., 104, 105
Asher, N., 2, 11, 13
Augustinova, M., 13
Aydede, M., 21

B

Baayen, R. H., 206, 307
Badre, David, 215
Baggio, Giosuè, 192
Baker, Mark C., 292
Baker, Shawn C., 216
Bard, E. G., 293
Baroni, M., 19, 20
Barsalou, L. W., 9–14, 16, 17, 19–22, 24, 39, 41, 105–107, 113, 225, 236
Bartsch, Renate, 295
Bates, D. M., 206, 230, 307
Batterman, N., 12
Battig, W. F., 106
Bereshpolova, Yulia, 217
Berlin, B., 123
Binder, J. R., 215
Blanchette, I., 12
Blok, S., 106
Bosch, P., 13
Bourne, L., 68
Bowerman, M., 100

Boyes-Braem, P., 114
Bransford, J. D., 22
Braun, D., 113
Braze, David, 191, 192, 194, 200, 202, 205, 206, 208, 214, 217
Breazeal, C., 11, 21
Brennan, R. E., 240
Brooks, L. R., 16
Bruni, E., 20
Bub, D. N., 19
Buhrmester, M., 304
Burge, T., 38
Burnett, R., 106

C

Caramazza, A., 12
Carlson, G., 137
Carlson, G. N., 137
Casasanto, D., 11, 13
Castellanos, F. Xavier, 216
Cavanna, Andrea E., 216
Cesario, J., 13
Chaigneau, S. E., 19
Chen, Lillian, 192
Chen, Yao, 217
Chierchia, G., 124, 128, 136
Choi, S., 100
Choma, Travis, 192
Chomsky, N., 43, 48, 49, 68, 69
Clark, A., 15, 21
Cohen, B., 2
Collins, Louis, 210
Conant, Lisa L., 215
Connell, L., 11, 13, 19
Connolly, A. C., 225, 227, 233
Conrad, C., 12, 13
Constantinescu, Camelia, 314
Cook, R. S., 127
Costantini, Marcello, 216

Costello, F. J., 9
 Cramon, D. Yves Von, 216
 Crozier, Sophie, 216
 Cruse, D. A., 139

D

Dai, Weiqian, 210
 Dalrymple, M., 141
 Damasio, Antonio R., 215
 Damasio, Hanna, 215
 Davidson, D. J., 206, 307
 David, Wattenmaker W., 299, 314
 Davies, Caroline, 102
 Davies, Mark, 305
 Davis, M. H., 150
 Deacon, T. W., 9
 De Boeck, P., 114
 Dennis, S., 19
 Deo, Ashwini, 191, 192, 194–196, 198, 200,
 202, 205, 206, 208, 214, 217
 De Ruiter, J. P., 126
 Desai, Rutvik H., 215
 De Swart, H., 195
 De Vries, Hanna, 314
 Dierckx, Rudi, 209
 Dillingham, E. M., 228, 237
 DiStasio, Marcello, 209
 Ditye, Thomas, 216
 Dolan, Raymond J., 216
 Donald, M., 9
 Dove, G., 12
 Drašković, I., 124, 128
 Dubois, Bruno, 216
 Dubois, D., 101, 102, 104, 107
 Dumais, S. T., 19
 Dunbar, K., 12

E

Erik, Nolfé, 209
 Erk, K., 19
 Espinal, M. Teresa., 314
 Estes, Z., 105, 113
 Evans, Alan C., 210
 Evans, V., 11, 13

F

Faltz, L. M., 124, 128, 140
 Feng, Ching-Mei, 217
 Ferrand, L., 13
 Feser, E., 236–238
 Fletcher, Paul C., 216
 Fodor, J. A., 2, 95, 104, 225, 227, 233, 280
 Forbus, K. D., 12
 Fox, Danny, 300

Fox, Peter T., 217
 Frackowiak, Richard S. J., 216
 Francez, Itamar, 196
 Frassinetti, Francesca, 216
 Frazier, Lyn, 201
 Frege, G., 38, 66–68, 70, 72, 83, 85, 245, 251
 Friedmann, N., 140, 155
 Frisson, Steven, 192
 Frith, Chris D., 216

G

Gagné, C. L., 5, 9, 17, 118, 224–229, 231–235,
 237, 238, 241, 242
 Gao, Jia-Hong, 217
 Gärdenfors, P., 136, 303
 Gardiner, M. M., 107, 109
 Gawron, Jean Mark, 196
 Gawronski, B., 13
 Geeraerts, D., 126
 Gentner, D., 12
 Giannakidou, Anastasia, 293
 Gibbert, M., 113
 Glaser, W. R., 19
 Gleitman, H., 104, 105, 225, 227, 233
 Gleitman, L. R., 104, 105, 225, 227, 233
 Glucksberg, S., 101, 102, 110, 118, 140, 297
 Goodenough, W. H., 101
 Gordon, Betty, 215
 Gosling, D., 304
 Grabowski, Thomas J., 215
 Grafman, Jordan, 216
 Graf, Markus, 216
 Graves, William W., 215
 Gray, W. D., 114
 Greenspan, S. L., 12, 13
 Grimshaw, Jane, 201
 Gurman, Ellen, 293

H

Hack, Nancy, 215
 Hackl, Martin, 300
 Hagoort, Peter, 192
 Hale, C. R., 16, 225
 Hampton, J. A., 15, 97–99, 101, 102, 104, 105,
 107, 108, 110, 112, 113, 124, 125, 128,
 135, 136, 144, 158, 225–228, 231, 233,
 293, 297, 298, 302, 304, 310
 Hancock, C., 238
 Hazen, A., 40
 Heard, A. W., 18, 22
 Heider, E. R., 123, 127
 Heim, I., 136, 292
 Heussen, Daniel, 298, 304, 310
 Heycock, C., 140

Hichwa, Richard D., 215
 Hickok, Gregory, 215
 Hintzen, W., 96
 Hoeksema, Jack, 300
 Hoge, Rick, 210
 Holmes, Colin J., 210
 Holyoak, K. J., 106
 Honda, Manabu, 216
 Horschig, J. M., 123
 Hummel, J. E., 106
 Humphries, Colin, 215
 Husband, E. Matthew, 193, 207, 217
 Husserl, E., 38

I

Insler, Rachel Z., 215

J

Jackendoff, Ray, 192, 196
 Jackowski, Marcel P., 209
 Jacobs, Filip, 209
 Jacobson, P., 139
 Ji, J., 225
 Johnson, D. M., 114
 Johnson, K., 96
 Johnson, M. K., 22
 Jones, M., 11
 Jönsson, M. L., 98, 109, 225–228, 231, 233

K

Kahneman, D., 107
 Kamp, H., 2, 136, 294
 Kanazawa, M., 141
 Katsika, Argyro, 191, 192, 194, 200, 202, 205, 206, 208, 214, 217
 Katz, J. J., 2
 Kay, P., 123, 127
 Keane, M. T., 9, 225
 Keenan, E. L., 124, 128, 140
 Keil, F. C., 12
 Keller, Frank, 192
 Kelly, Lisa A., 193, 207, 217
 Kennedy, C., 2, 136, 292, 294
 Kerem, N., 140, 155
 Khemlani, S., 140
 Killackey, Jennifer, 215
 Kim, M., 141
 Kim, S., 106
 Kintsch, W., 19
 Kittur, A., 106
 Klein, E., 136, 294
 Koontz-Garboden, Andrew, 196
 Kostelecky, M., 240
 Krain, Amy L., 216

Kratzer, A., 136
 Krifka, M., 142, 195
 Kripke, S., 38
 Kwang, T., 304

L

Lakoff, G., 101
 Lambalgen, Michiel Van, 192
 Landauer, T. K., 19
 Landman, Fred, 295
 Langacker, R., 19, 24
 Lapata, M., 19, 192
 Larochelle, S., 105
 Lasersohn, Peter, 293
 Lau, Ellen F., 215
 Laurence, S., 225, 234
 LeBihan, Denis, 216
 Lebois, L. A. M., 11, 14
 Lee, C., 158
 Lehéricy, Stéphane, 216
 Lemahieu, Ignace, 209
 Lenci, A., 19, 194, 206, 208, 214
 Lepore, E., 95, 233
 Leslie, S., 140
 Levin, Beth, 217
 Lewis, D., 114
 Liebenthal, Einat, 215
 Link, G., 140
 Longe, Olivia A., 216
 Lopopolo, M., 20
 Louwerse, M. M., 19
 Love, B. C., 11
 Lupyan, G., 11, 13
 Lynott, D., 11, 13

M

Machery, E., 10, 12, 96
 Macknik, Stephen L., 217
 Madden, J. D., 238
 Magnani, Barbara, 216
 Mahon, B. Z., 12
 Majid, A., 123
 Malt, B. C., 105
 Margolis, E., 225, 234
 Marr, D., 24
 Marrett, Sean, 210
 Martin, Andrea E., 193
 Martinez-Conde, Susana, 217
 Martin-Löf, P., 54, 250
 Masson, M. E., 19
 Mathôt, S., 150
 Mazursky, D., 113
 McClelland, J. L., 11, 15, 16
 McCloskey, M. E., 13, 101, 297

- McConnell-Ginet, S., 124, 128, 136
 McElree, Brian, 192, 193, 207, 214, 215
 Mchombo, S., 141
 McNally, L., 2, 136
 McNamara, D. S., 19
 Medin, D. L., 9, 16, 18, 106, 117, 123, 124, 298
 Medler, David A., 215
 Mercier, D.-J., 234
 Mervis, C. B., 114, 144, 298
 Meyer, Ernst, 210
 Michaelis, Laura A., 214
 Milham, Michael P., 216
 Milot, Sylvain, 210
 Minda, J. P., 11, 16
 Mirza, H., 110
 Mitterer, H., 123, 126, 127
 Mohr, Sibylle, 192
 Moltmann, F., 300
 Montague, R., 2, 106
 Morzycki, Marcin, 293, 313
 Mouridsen, Kim, 216
 Mullaly, A. C., 225
 Murphy, G. L., 2, 100, 106, 117, 225, 297, 298
 Müsseler, J., 123
- N**
- Neelin, Peter, 210
 Nouwen, R. R., 2
- O**
- Okuda, Hirohito, 209
 Oliveri, Massimiliano, 216
 Osherson, D. N., 2, 5, 33, 96, 98, 104, 105, 108, 109, 123–125, 135, 136, 225
- P**
- Padó, S., 19, 194, 206, 208, 214
 Paivio, A., 19
 Papademetris, Xenophon, 209
 Papies, E. K., 13
 Paré-Blagojev, E. Juliana, 215
 Parmar, S., 110
 Partee, B., 2, 140, 192, 136
 Passanisi, A., 225–228, 231, 233
 Perler, D., 236
 Pesetsky, David, 217
 Peters, S., 141
 Phillips, Colin, 215
 Philip, W., 143
 Pickering, Martin J., 192, 193, 214, 215
 Pillon, Bernard, 216
 Piñango, Maria Mercedes, 191, 192, 194–196, 198, 202, 205–208, 214, 217
- Pinheiro, J. C., 230
 Poeppel, David, 215
 Poesio, M., 20
 Poldrack, Russell A., 215
 Poortman, E. B., 155
 Pothos, E. M., 11, 16
 Prasada, S., 228, 237
 Price, Cathy J., 217
 Prinz, Wolfgang, 216
 Pullum, G. K., 96
 Pustejovsky, J., 2, 124, 192, 202, 238
 Putnam, H., 38
 Pykkänen, Liina, 192, 193, 207
- Q**
- Quené, H., 230
- R**
- Rajeevan, Nallakkandi, 209
 Randall, Billi, 216
 Rattermann, M. J., 12
 Rayner, Keith, 201
 R Core Team, 206
 Regier, T., 127
 Rey, G., 105
 Reynolds, P. L., 234
 Rice, S. A., 100
 Richard, S., 105
 Rips, L. J., 105, 144, 225
 Robbins, P., 21
 Robertson, Dan, 293
 Rodd, Jennifer M., 216
 Roepstorff, Andreas, 216
 Rooth, M., 140, 192
 Rosch, E., 2, 106, 109, 114, 144, 297, 298
 Ross, N., 106
 Rosseel, Y., 114
 Roth, Emilie, 297
 Rumelhart, D. E., 11, 15, 16
 Ruts, W., 98
- S**
- Sadato, Norihiro, 216
 Santos, A., 19
 Sassoan, Galit W., 137, 295, 300–303, 312
 Schaffer, Marguerite M., 298
 Schank, R., 61
 Scheeper, Christoph, 192
 Scholz, B. C., 96
 Schreij, D., 150
 Schreuder, R., 124
 Schubotz, Ricarda I., 216
 Sedivy, J., 136
 Seely, Rachel E., 192

Shallice, Tim, 216
 Shamir, Adi, 300, 301, 311, 312
 Shapiro, Lewis P., 201, 215
 Shoben, E. J., 9, 12, 18, 105, 118, 123, 124, 144, 225, 297
 Simmons, Claire L., 298, 304, 310
 Simmons, S., 105
 Simmons, W. K., 10, 19
 Sirigu, Angela, 216
 Sloman, S. A., 19, 105
 Slovic, P., 107
 Smart, Andrew, 193
 Smith, E. E., 98, 104, 105, 108, 123–125, 135, 136, 144, 225
 Smith, J. D., 11, 16
 Smith, L. B., 2, 5, 9, 11, 12, 21
 Smits, T., 114
 Sorace, Antonella, 293
 Soulières, I., 105
 Spalding, T. L., 5, 9, 17, 224–229, 231–235, 237, 238, 241, 242
 Speelman, D., 126
 Springer, Anne, 216
 Stadler, Waltraud, 216
 Staib, Lawrence H., 209
 Stedman, J., 238
 Steinvall, A., 126
 Østergaard, Leif, 216
 Storms, Gert, 98, 114, 298, 304, 310
 Struiksma, M. E., 155
 Swadlow, Harvey A., 217

T

Tan, Li Hai, 217
 Tanaka, Satoshi, 216
 Theeuwes, J., 150
 Todd Constable, R., 209
 Toga, Arthur W., 210
 Toledo, A., 137
 Tranel, Daniel, 215
 Traxler, Matthew J., 192, 193, 214, 215
 Trimble, Michael R., 216
 Tversky, A., 107, 298
 Tyler, Lorraine K., 216

U

Utt, Jason, 194, 206, 208, 214

V

Van Casteren, M., 150
 Van de Moortele, Pierre-François, 216
 Van den Bergh, H., 230
 Vandenbroucke, A., 98
 Verkuyl, Henk, 196
 Voet, Tony, 209
 Von Fintel, Kai, 300
 Von Stechow, Armin, 292

W

Wagner, Anthony D., 215
 Wallentin, Mikkel, 216
 Walsh, Vincent, 216
 Weed, Ethan, 216
 Wernicke's area, 192
 Werning, M., 96
 Wiemer-Hastings, K., 13
 Williams, Rihana S., 192, 214
 Wills, A. J., 11, 16
 Wilson, Amanda M., 216
 Wilson, C. D., 19
 Wilson-Mendenhall, C. D., 11
 Winter, Y., 1, 4, 13, 100, 106, 112, 137, 140, 142, 143, 154, 155, 158, 159
 Wippel, J. F., 234
 Wisniewski, E. J., 118, 225
 Wittgenstein, L., 297
 Woods, Roger, 210
 Worsley, Keith, 210
 Wright, C., 108
 Wu, L., 236

Y

Yeh, W., 101, 102, 104, 107, 108
 Yoon, Suwon, 293

Z

Zamparelli, R., 141
 Zarcone, Alessandra, 194, 206, 208, 214
 Zhu, David C., 193, 207, 217
 Zurif, Edgar B., 191, 192, 201, 207, 217
 Zwarts, Joost, 196

Subject Index

A

- Ability vs. activity, 80
- Acceptability judgments, 134, 166, 182, 303
- Accomplishment (verbs), 80
- Achievement (verbs), 80
- Activity (verbs), 4, 80, 102, 207, 216, 238, 271, 282
- Additive similarity, 298, 299, 313
- Adjective, 6, 100, 102, 108, 124, 125, 131, 133, 134, 136–138, 144, 164, 168, 172, 173, 177, 246, 248–250, 253–256, 260–263, 271, 272, 274, 279, 284, 291, 292, 294, 295, 299–301, 309–313, 315
- Adjective-noun combination, 123–125, 127, 131, 133, 134, 263
- Affordance, 5, 19, 245, 246, 251, 252, 256, 259
- Ambiguity, 100, 107, 141, 164, 165, 191, 200, 207, 214–216
- Ambiguity resolution, 201 *See also* Ambiguity
- Argument saturation, 271
- Aristotelian-thomistic approach, 233
- Aspectual verbs, 191, 193–195, 197, 199, 200, 202, 203, 206, 207, 210, 214–216, 218, 219
- Atomism, 32, 33, 51
- Atomistic (theory of meaning), 61, 70, 86
- Axis, 191, 196–200, 216

B

- Basic level (of categorization), 113, 114, 116
- Behaviourism, 43
- Belief-Desire-Action, 84
- Biological classification, 298
- Boolean, 136, 140
- Brain in a Vat, 59
- Broca's area, 215, 217
- Brodman area, 193, 207, 210, 212, 213, 215, 217
- Building-block compositionality, 35

C

- CARIN model, 276
- Categorization, 98, 103, 104, 107, 109, 111, 123, 125, 127, 128, 132, 134, 136, 173, 237, 273, 292, 293, 295–297, 299, 302, 304, 310, 311, 314
- Category membership, 12, 97, 106, 110, 129, 144, 148, 153, 169, 172
- Character, 11, 66, 129, 223, 250
- Circumstance of utterance, 66
- Classical theory (of concepts), 105
- Classification, 103, 107, 126, 293, 294, 298, 299
- Coercion, 73, 74, 191, 193–195, 199, 200, 202, 207, 214, 215, 217, 219, 250
- Coercion configuration, 199 *See also* Complement coercion
- Coercion verb, 192 *See also* Complement coercion
- Cognitive linguistics, 19, 45–47, 50, 56, 73 *See also* Emergentism
- Color adjectives, 123, 126, 127, 133, 136, 173
- Combined concepts, 20, 223, 224, 227, 239, 253, 281
- Communication, 15, 209, 232, 234
- Comparison constructions, 294, 302, 310, 314
- Complement coercion, 192, 194, 200, 208
- Compositionality, 2, 3, 95, 96, 112, 163, 164, 173, 176, 187
- Compositional semantics, 140, 158, 163
- Compound words, 224
- Concept, 1–6, 10, 11, 13–24, 96, 98, 100, 101, 106, 109, 112, 114, 116, 117, 124, 125, 136, 143–146, 154, 156–158, 164–167, 169–171, 173–187, 197, 223, 225–227, 232–237, 240, 242, 245, 247, 250–256, 259–263, 272–274, 276, 278–281, 283, 296, 298, 299, 312

- Concept composition, 1, 2, 4, 5, 12, 13, 16–21, 23–25, 124, 127, 173, 177, 219, 245–247, 252, 256, 261, 264
 Conceptual cores, 12
 Conceptualization, 216, 240
 Conceptual role, 62
 Conceptual specificity, 270, 272, 274, 281, 283
 Conceptual structure, 10, 105, 155, 251, 263, 277, 297
 Conditional, 165, 196
 Conjunction, 3, 97–99, 105, 109, 112, 114, 141–143, 145, 154–156, 158, 171, 179
 Connectionist, 11
 Canonical derivability conditions, 54
 Construction grammar, 2
 Content, 50, 53, 57, 64, 66, 70, 71, 248, 250
 Content variability, 9, 11, 12
 Context, 3, 101–103, 108, 116, 124, 127, 143, 157, 164, 166, 173, 187, 198, 240, 281, 292, 299, 301, 312
 Context dependence, 12, 13, 313
 Co-occurrence (in theories of meaning), 43
 Core (of concept), 12
 Creative, 3, 60, 113, 116, 117
 Creativity, 60, 87, 113
 Critical typicality, 105, 163, 175–181, 183, 185
- D**
- Default interpretation, 19, 246, 250, 252, 264, 310
 Definitional (aspects of meaning), 143
 Degree, 2, 6, 97, 104–106, 108, 109, 123, 125, 136, 153, 155, 165, 169, 218, 228, 292–296, 298, 310, 313
 Dictionary, 116, 126
 Dimension, 105, 108, 125, 133, 166, 197, 198, 201, 202, 210, 216, 275, 278, 280, 284, 294, 296, 298, 300–302, 313, 314
 Dimension function, 200, 215, 216
 Dimension integration, 298
 Directed Path Structure (DPS), 195
 Discourse Representation Structure (DRS), 260, 261
 Discourse Representation Theory (DRT), 259–261, 263
 Disjunction, 3, 98, 112, 172
 Disjunctive concepts, 112
 Distributional semantics, 5, 19, 175, 245, 247, 250, 252, 256, 257, 259, 260, 263
 Distributivity operator, 140
 Dual-aspect (lexical items), 54, 74
- E**
- Embodied concepts. *See* Grounding (of concepts)
 Emergentism, 42, 46, 48
 Emergent (property), 3 *See also* Emergentism
 Encyclopedic (knowledge), 40, 47, 55, 62, 74, 76, 86
 Entry rule, 54
 Ergative (verbs), 80
 Essentialism, 237
 Evaluation rule, 54
 Eventive (verbs), 193
 Evolution, 2, 9, 299
 Exception phrases, 295, 300–302, 304, 310–312
 Exhaustive activation, 200, 201, 210, 215
 Exit rule, 54
 Extension, 96, 97, 100–102, 113, 129, 163, 169, 176, 251, 291, 297
 Extensionality, 37
 Externalist (approach to meaning), 3, 5, 43, 116
- F**
- First-pass composition, 16
 fMRI, 4, 202, 207, 209, 212, 215–217
 Focal (color), 123, 127, 131, 166, 170, 174–176, 182, 183
 Formal semantics, 164
 Functional compositionality, 171
- G**
- General terms, 51, 57, 146, 177, 184
 Generative Lexicon, 256
 Generic (statements), 5
 Gestaltism, 33
 Gestaltism. *See* Gestalt
 Gradability, 6, 165, 291–293, 302, 313, 314
 Gradable adjectives, 137, 165, 171, 292, 314
 Graded concept, 144, 175, 182
 Gradedness, 104
 Grasp, 51, 66, 70
 Grounded cognition, 25
 Grounding (of concepts), 59
 Guppy effects, 172, 177
- H**
- Historical-causal theory (of names), 76
 Holism, 36, 63, 169
 Holistic theory of meaning, 61, 86 *See also* Holism
 Homography, 56

Homophony, 56
 Human cognition, 25, 224, 238

I

Ideals, 11, 105, 106, 118
 Incremental theme, 200
 Indexicality, 248
 Individual differences (in belief), 15, 98
 Inferentialism, 43, 54
 Informational significance, 50
 Information-theory, 15
 Instantiation, 106, 113, 114, 116, 137, 245
 Intension, 100, 102, 251, 291
 Intersection, 97, 123, 127, 136, 140, 142, 186, 262
 Intersective (conceptual combination), 118
 Intuitionistic logic, 250
 Inverse thematic function, 198

K

Kind, 2, 5, 10, 13, 17, 20, 21, 23, 101, 104, 110, 112, 117, 129, 133, 136, 168, 169, 172, 175, 180, 216, 225, 235, 240, 257, 258, 260, 261, 299, 312

L

Lambda term, 73
 Language game, 43
 Language of thought, 1
 Learnability, 60, 87
 Left anterior temporal lobe, 5
 Left Inferior Frontal cortex (LIF cortex), 213, 215
 Lexical information, 139–142, 158 *See also* Lexical knowledge
 Lexical item, 116, 215, 238
 Lexicalization, 219
 Lexical knowledge, 252
 Lexical semantics, 139, 193, 194, 203
 Literalism, 46–49
 Logical expression, 139, 158
 Logical semantics, 139

M

Magnetoencephalography, 5, 271
 Maximality principle, 142
 Maximal Typicality Hypothesis (MTH), 144, 154, 178
 Meaning composition, 2, 124, 136, 144, 164, 177, 217, 219
 Meaning function, 34
 Membership judgment, 67
 Mental item, 40

Mental representation, 1, 3, 124
 Mental state, 9
 Metalinguistic, 293
 Metaphorical interpretation, 77
 Metonymy, 199
 Minimalist program, 68
 Mode of presentation, 66–68
 Modification, 5, 21, 100, 124, 126, 136, 172, 175, 233, 234, 245, 248–250, 253–255, 261–263, 271, 274, 275, 277, 281, 283, 284, 301, 311
 Modification effect, 225–228, 230, 231, 233, 234, 239, 240, 242, 281
 Morphological gradability, 291, 293, 314
See also ‘Degree morpheme’ and ‘gradable’
 MTH. *See* Maximal Typicality Hypothesis
 Multidimensional, 294, 295, 301, 302, 311
 Multiplicative similarity, 298, 299

N

Narrow content, 64, 70
 Narrow meaning, 64 *See also* Narrow content
 Natural deduction, 54
 Natural kind, 4, 126, 260, 299, 310, 312
 Negation, 25, 109, 112, 172, 300, 301, 310
 Neo-empiricist (theories of concepts), 41
 Nexus, 64, 65, 68–71, 87
 Nominal compounding, 223
 Non-intersective (conceptual combination), 139
 Non-intersective conjunction, 142 *See also* Intersective conjunction
 Noun, 3, 5, 18, 96, 108, 114–116, 118, 124, 125, 129, 131, 133, 136, 137, 142, 144, 155, 157, 223, 225, 229, 231, 242, 271–274, 276, 279, 280, 282, 284, 291–293, 295, 296, 298, 301, 302, 307, 310–314
 Noun phrase conjunction, 142
 Numeral quantification, 270

O

Objectivist, 247
 Ontological compositionality, 32–35, 42, 46
 Operator, 140, 249, 250
 Ordinary language, 54
 Overextension, 97–100, 109, 117, 125, 136, 163, 173

P

Performance, 304
 Periphery (of concept), 40

- Plurality, 163, 165, 175, 177, 179, 186, 187
 Plural predicate conjunction, 140, 142, 147, 149
 Plural predication, 142
 Plurals, 155, 156, 163, 165, 175
 Pragmatics, 2, 109
 Precuneus, 215
 Predicate transfer, 73
 Presupposition, 109, 197, 199
 Primitives (semantic relations), 248
 Private concepts, 51
 Productivity, 224
 Progressive (aspect), 80
 Proof theory, 250
 Proper name, 156
 Property verification, 224, 226, 231, 233, 234, 238–240
 Proposition, 238, 240, 293
 Prototype, 3, 5, 11, 16, 95, 96, 98, 100, 104, 105, 109, 112, 117, 123, 169, 182, 237, 291, 295, 297, 302
 Psychological plausibility, 9, 264
 Psychological verbs, 4, 191, 194, 200, 203, 204, 206, 214, 218, 219
 Psychologism, 38
 Public concepts, 85
 Public meaning, 51
- Q**
 Qualia, 238, 248, 250
 Quantification, 25, 291, 296, 299–301, 309–311
 Quantifier, 167, 177
- R**
 Real-time composition, 219
 Real-time comprehension, 192, 195, 214
 Reciprocals, 141–143
 Reference, 6, 126, 245, 251, 254, 259, 261, 264, 291
 Reference shift, 155–157
 Relation-based models, 269, 275–278, 281–284
 Representation, 1–3, 11, 13, 16–20, 23, 24, 103, 112, 124, 192, 193, 216, 226, 234, 235, 247, 249, 250, 252, 255–260, 263, 273, 274, 278–281, 283, 299
 Reverse modification effect, 280, 281
 RICE model, 277, 282
- S**
 Schema-based models, 269, 275, 277, 278, 281, 283
 Self-paced reading, 4, 191, 202, 204, 208, 214
 Semantic dementia, 272
 Semantic hub model, 273, 274, 278, 283
 Semantic liaison, 62
 Semantic memory, 14, 273
 Sense, 15, 100, 108, 129, 137, 166, 235, 238–240, 245, 251, 270, 272, 305, 314
 Sense transfer, 72
 Similarity, 19, 104, 105, 107, 124, 125, 169, 241, 277, 291, 295–298, 301–303, 312
 Situated cognition, 23
 Situated (concept), 11 *See also* Grounding
 Size adjectives, 313, 314
 Sociolinguistics, 47
 Specificity (of meaning), 85
 Stative (verbs), 80
 Strongest Meaning Hypothesis (SMH), 141, 177
 Structured individual analysis, 198–200, 214, 219
 Structured individual hypothesis, 200, 202, 208, 215, 217, 219
 Subjectivism, 11, 247
- T**
 Truth-functional (theory), 50
 Truth theory, 3
 Truth-value, 4, 149, 163, 165, 166, 173, 178, 187
 Two-tier (theory), 51, 54, 55, 64, 68, 70, 88
 Type, 22, 23, 96, 126, 133, 150, 191, 192, 196, 202, 224, 226, 229, 231, 239, 250, 257, 269, 270, 277, 278, 292, 300, 306–309, 312, 314
 Type coercion, 192 *See also* coercion
 Type-shifting, 192 *See also* Coercion, Type coercion
 Typicality, 3, 4, 97, 103–110, 118, 123–128, 130, 134, 135, 141, 143–147, 152, 154–158, 163, 165, 172–177, 180–182, 187, 226, 256, 257, 279, 293, 297
 Typicality judgment, 133 *See also* Typicality
- U**
 Underspecification, 199
 Ungrammaticality, 78
 Universal quantification, 142, 301, 311
- V**
 Vagueness, 2, 102, 103, 107, 108, 110, 111, 144, 165–167, 169, 171, 172, 177, 187
 Vector, 191, 207, 213, 215, 218, 256–258, 260–262

W

Wernicke's area, [191](#), [192](#), [207](#), [211](#), [213](#), [215](#),
[217](#), [218](#)
Wholism, [33](#), [36](#), [61](#)

Word meaning, [5](#), [96](#), [127](#), [140](#), [257](#)
Word order, [269](#), [272](#), [273](#)
World knowledge, [113](#), [123–126](#), [128](#), [133](#),
[137](#), [276](#), [277](#), [282](#), [283](#)

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

