Explain Me This

Creativity, Competition, and the Partial Productivity of Constructions

Adele E. Goldberg

Princeton University Press Princeton and Oxford

Published by Princeton University Press
41 William Street, Princeton, New Jersey 08540
6 Oxford Street, Woodstock, Oxfordshire OX20 1TR

press.princeton.edu

Library of Congress Control Number: (to come)
ISBN 978-0-691-17425-9
ISBN (pbk.) 978-0691-17426-6

British Library Cataloging-in-Publication Data is available

Editorial: Meagan Levinson and Samantha Nader
Production Editorial: Jenny Wolkowicki
Jacket/Cover Credit: {~?~to come}
Production: Erin Suydam
Publicity:
Copyeditor: Maia Vaswani

This book has been composed in Adobe Text Pro and Gotham

Printed on acid-free paper. ∞

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Chapter 1

Introduction

It is easy to take our knowledge of language for granted. We learn language before we carry our first backpack to school, and we use it almost every waking hour of every day. Although we may not have studied quantum theory, or read Homer or James Joyce, we are each expert at using our own native language. The challenge that all learners face becomes more apparent when we try to learn a second language in school or as adults.

There are many utterances that are perfectly understandable, but which nonetheless tend to be avoided by native speakers of English. If asked, speakers will agree that there is something mildly "off" about them, even though they may have difficulty articulating exactly why they don't sound quite right. For example, we might confess that someone is driving us *crazy* (or *bananas* or *insane*), but we know that it would sound odd to complain that someone is driving us *angry*. We know that *tall* bushes are *high* bushes, but a *high* teenager is not necessarily *tall.* We can be creative in how language is used, but our creativity is constrained in ways that can be hard to articulate. For example, someone can *tell me something* or *tell something to me*, but they can only *explain this to me*; that is, it sounds somewhat unconventional to native speakers of English to say, *explain me this*. That is what this book aims to explain: when, why, and how native speakers are sometimes creative with language and yet at other times much more conservative.

Speakers avoid saying certain things, of course, simply because they want to avoid overtly negative reactions. The following are examples of such ill-advised utterances:

Sorry Mom, I didn't mean to get caught.
I only care about my grade in this course.
Your nose is too big for your face.

But children are not systematically corrected for the types of utterances this volume aims to address, which will hereafter be indicated by a preceding "?" (?*explain me this*, ?*drive him angry*, etc.). Caregivers are much more focused on the content of children's speech than on its form, as long as the message is clear enough. For example, a child who says *Me loves you, mommy* is more likely to

get a hug than a grammar lesson, and a young child who utters an impressively grammatical utterance such as *I have just completed a mural on the living-room wall with indelible markers* is unlikely to get positive feedback from most parents. The sorts of formulations that native speakers recognize as odd are also not the sorts of formulations that grammar teachers warn against, since they are so rarely uttered by native speakers that no admonishment is needed.

To be clear, it is not that one *never* hears expressions such as ?*explain me this* (or ?*drive him angry*), or that all speakers judge them to be equally odd. In fact, speakers' judgments are gradient and dependent on a number of interrelated factors that are the focus of this book. But corpus and experimental studies confirm that certain types of utterances are avoided by native speakers much more than would be expected by chance. In order to think about how these aspects of language are learned, it's worth thinking about what speakers and language learners are trying to do.

1.1 The Puzzle

The learner's goal is to comprehend messages, given the forms she witnesses, and to produce forms, given the messages she wants to convey. Therefore, speakers must learn the ways in which forms and functions are paired in the language(s) they speak. These learned pairings of forms and functions are referred to here as grammatical CONSTRUCTIONS. Speakers also aim to express their intended messages efficiently and effectively while respecting the conventions of their speech communities, as discussed more below.

Constructions generally allow us to apply our linguistic knowledge to new situations and experiences. English tends to be particularly flexible in the ways in which constructions are PRODUCTIVE. A few examples of productive uses of familiar constructions are provided in table 1.1, with labels for each grammatical construction provided on the right.

TABLE 1.1. Novel linguistic exemplars that demonstrate the productivity of various constructions

"Hey man, *bust me some fries*."	Double-object construction
"Can we *vulture your table*?"	Transitive causative construction
"Vernon *tweeted to say she doesn't like us*."	*To* infinitive construction
"What a *bodacious thing* to say."	Attributive modification construction

Attested examples are cited in quotation marks. Here and below unless otherwise noted, attested examples come from Google.

TABLE 1.2. Novel formulations that are judged odd by native speakers

?She explained him the story. (cf. She told/guaranteed him the story.)	Double-object construction
?He vanished the rabbit. (cf. He hid/banished the rabbit.)	Transitive causative construction
?She considered to say something. (cf. She hoped/planned to say something.)	*To* infinitive construction
?The asleep boy (cf. The astute/sleepy boy)	Attributive modification construction

At the same time, the constructions exemplified in table 1.1 resist being used productively with certain verbs or adjectives, even when the intended meaning is perfectly clear. Examples that illustrate the lack of full productivity are provided in table 1.2. Under each ill-formed example is a closely related fully acceptable example, in parentheses. The latter are provided to indicate that there are no simple, system-wide explanations for why the odd sentences strike native speakers of English as odd. Thus, constructions can be extended for use with some words (table 1.1), but they are rarely completely productive (table 1.2), even when no general constraints are violated. How is it that native speakers know to avoid certain expressions while nonetheless using language in creative ways? It is no exaggeration to say that this basic question has bedeviled linguists and psychologists for the past four decades.

1.2 The Roadmap

The paradox of PARTIAL PRODUCTIVITY of constructions is what this book aims to address. We will also address several issues that have not widely been viewed as directly related. In particular, chapter 2 includes a discussion of how we learn to circumscribe the meanings of words. Close attention to word meanings reveals that speakers possess a vast amount of rich contextual knowledge about what each word means, and about which other words it tends to co-occur with. But, initially, young children make certain errors. They may call the moon a *ball*, or the mailman *Daddy*, before they learn and become fluent with other words (specifically, *moon* and *mailman*). That is, children need to learn to restrict their use of individual words by witnessing how those words *and other words* are used in particular contexts. The rest of the book argues that the same mechanisms involved in learning and restricting word meanings are used when learning and restricting grammatical constructions, and that

this process explains how we come to avoid formulations such as ?*explain me this*. By beginning with word meanings, I hope to make the discussion of our primary target—the partial productivity of grammatical constructions—more accessible. That is, once we have a better understanding of word meanings, we can tackle grammatical constructions by essentially asking: What would words do?

Chapter 3 outlines the various factors that are relevant to our knowledge of how grammatical constructions are used within a given speech community. These include formal properties (*syntax*), words and partially filled words (*morphology*), meaning (*semantics*), discourse function (*information structure*), and social context. An appreciation of these factors is a prerequisite for solving the *explain-me-this* puzzle. This chapter also highlights the remarkable degree of cross-linguistic variation that exists in how simple clauses are expressed in the world's languages, in an effort to emphasize just how much people must learn in order to use the constructions in their language appropriately.

The proposed solution to the partial productivity puzzle allows both generalizations (table 1.1) and exceptions (table 1.2) to be learned via the same mechanisms. In particular, in chapters 4 and 5, two key factors—COVERAGE and COMPETITION—are discussed. Chapter 4 explains how constraints on meaning and use *emerge*, as witnessed exemplars cluster within the high-dimensional conceptual space in which our representations for language exist. This chapter outlines how clustering licenses creative uses of constructions. In particular, a single factor, COVERAGE, combines variability, type frequency, and similarity; specifically, a new instance is licensed to the extent that the ad hoc category required to contain it has been well attested (has been sufficiently "covered"). Also outlined in this chapter is a useful model for formalizing the required mechanism; namely, an incremental Bayesian clustering algorithm (Barak et al., 2014, 2016; see also Alishahi and Stevenson, 2008; Matusevych et al., 2017).

In chapter 5, the critical role of competition is detailed. As we comprehend utterances, we attempt to anticipate what the speaker will say next, and we are able to use what the speaker actually says to improve future predictions through a process of error-driven learning. Repeatedly witnessing certain formulations in certain types of contexts strengthens the connections between those grammatical constructions and the intended messages-in-context expressed; this results in conventional formulations becoming more accessible for expressing the types of messages that have been previously witnessed. When there exists a readily available formulation that expresses the intended message in the given context, it usually wins out over potential novel formulations. A special effort is required to buck conventional formulations, although this is possible, for the sake of memorability or playfulness (as in the title of this book). But when there is no

readily accessible combination of constructions available to express a speaker's intended message-in-context, she needs to extend language creatively.

The proposal is situated in a larger context in chapter 6. Many studies have demonstrated that children are initially less creative than adults: children behave "conservatively" in that they generalize constructions less freely than adults do. Yet other studies have found that children generalize *more* broadly than adults. This apparent paradox is reconciled by recognizing that children are less adept at aligning bits of knowledge within their high-dimensional conceptual space: sometimes they fail to recognize relevant parallels across exemplars, at least with sufficient confidence (and so they behave conservatively); other times they fail to recognize or retain relevant distinctions (and so they generalize or simplify). Appropriate use of grammatical constructions emerges once the relevant conditioning factors for each construction are learned, and the language user becomes more fluent at accessing the appropriate constructions from memory.

Chapter 6 also outlines why adult learners of a second language tend to have particular difficulty avoiding the types of odd formulations this book addresses (including ?*explain me this*). The suggested reasons go beyond the fact that adults receive less input overall, and that the input they do receive is less well suited to learning. In particular, adult learners need to inhibit their well-practiced native language in order to process a new language, and this appears to lead to a reduced ability to take full advantage of the competition among constructions within the new language. Since competition is argued to be key to constraining generalizations via statistical preemption (chapter 5), second-language learners tend to be more vulnerable to producing certain types of formulations that make sense but which native speakers systematically avoid. Additionally, while adults are generally quicker to discern which dimensions of similarity and dissimilarity are relevant to clustering linguistic representations within their hyper-dimensional conceptual space, they are at the same time prone to miss very subtle similarities and distinctions that are not relevant in their first language.

1.3 The CENCE ME Principles

The basic understanding of language that this book outlines is based on the key ideas listed in table 1.3, which are discussed in detail in the following chapters. An acronym of the key words in these principles is EEMCNCE, but EEMCNCE would be impossible to pronounce. So, let us instead use an anagram of EEMCNCE: **CENCE ME**. "CENCE ME," pronounced "sense me," is intended to emphasize the importance of sensible communication. CENCE ME also usefully illustrates productivity, since the phrase itself is a novel use of the transitive construction. The CENCE ME principles spell out some key assumptions of the more general

TABLE 1.3. The CENCE ME principles

A. Speakers balance the need to be **E**xpressive and **E**fficient while conforming to the conventions of their speech communities.
B. Our **M**emory is vast but imperfect: memory traces are retained but partially abstract ("lossy").[a]
C. Lossy memories are aligned when they share relevant aspects of form and function, resulting in overlapping, emergent clusters of representations: **C**onstructions.
D. **N**ew information is related to old information, resulting in a rich network of constructions.
E. During production, multiple constructions are activated and **C**ompete with one another to express our intended message.
F. During comprehension, mismatches between what is expected and what is witnessed fine-tune our network of learned constructions via **E**rror-driven learning.

[a] Our representations are "lossy," a term from computer science, in the sense that they are not fully specified in all detail.

USAGE-BASED CONSTRUCTIONIST APPROACH to language that are widely shared (see, e.g., Bybee, 2010; Christiansen and Chater, 2016; Goldberg, 2006; Kapatsinski, 2018; Langacker, 1988; Tomasello, 2003; Traugott and Trousdale, 2013). The approach also shares much with memory-based EXEMPLAR-BASED MODELS (Aha et al., 1991; Bod, 2009; Bybee, 2002; Daelemans and van den Bosch, 2005; Gahl and Yu, 2006; Kruschke, 1992; Nosofsky, 1986). The CENCE ME approach emphasizes that exemplars—structured representations—cluster within a hyper-dimensional conceptual space giving rise to emergent constructions, which are then extendable as needed for the purpose of communication.

Individual languages can and do vary in striking ways, as will be emphasized, but the usage-based constructionist approach adopted here suggests that the CENCE ME principles are at work in every natural language, serving to constrain and shape the range of possible human languages. The present book emphasizes examples in English because the majority of the experimental and modeling work to be described has been done on English, and because English is the language I know best.

My understanding of what a construction is has evolved. Early on, I adopted the following definition:

> C is a CONSTRUCTION if and only if C is a form-meaning pair $<F_i, S_i>$ such that some aspect of F_i or some aspect of S_i is not strictly predictable from

> C's component parts or from other previously established constructions. (Goldberg, 1995, 4)

Later, I recognized that this definition was too narrow. Our knowledge of language comprises a network of constructions, and we clearly know and remember conventional expressions even if they are in no way idiosyncratic. So I broadened my definition of constructions as follows:

> Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency. (Goldberg, 2006, 5).

The present volume offers a still more inclusive understanding of what constructions are, motivated by a better appreciation of human memory, learning, and categorization. Here, as explained in the following chapters, constructions are understood to be emergent clusters of lossy memory traces that are aligned within our high- (hyper!) dimensional conceptual space on the basis of shared form, function, and contextual dimensions.

Proponents of alternative perspectives or readers who wish to compare the present proposal with other proposals in more detail may find chapter 7 particularly relevant. There, several recent alternative proposals that aim to account for the partial productivity of constructions are discussed. These include, for example, the idea that speakers avoid straying from what they have witnessed ("conservatism via entrenchment"), that it is useful to posit invisible syntactic diacritics or underlying structures without specifying how these are to be identified by learners, that putting a cap on the number of exceptions and a floor on the number of instances that follow a generalization will ensure how and when generalizations are productive (the Tolerance and Sufficiency principles of Yang [2016]), or that incorporating degrees of uncertainty into formal rules is predictive (O'Donnell, 2015). While aspects of each of these proposals have merit, we will see that the usage-based constructionist approach, described by the CENCE ME principles, explains the facts more fully. The final chapter stands back and puts the discussion in a broader context, while raising several outstanding issues that remain to be addressed.

1.4 Speakers Are Efficient and Expressive and also Conform

Before leaving this introductory chapter, let's go over the first of the CENCE ME principles: We aim to express our messages effectively and efficiently while

obeying the conventions of our speech communities. To clarify the key terms involved:

1. **Expressiveness**: Linguistic options must be sufficient for conveying speaker's thoughts, beliefs, and attitudes in ways that listeners are able to understand.
2. **Efficiency**: Fewer and shorter constructions are easier to learn and produce than more or longer constructions.
3. **Obeying conventions**: Learners attempt to use language in the ways that others in their language communities do.

A language is only sufficiently expressive if it has the means to adequately convey a speaker's thoughts, beliefs, and attitudes in ways that avoid failures of communication. A *maximally expressive* language might have an ever-increasing number of words and constructions, with every potential distinction indicated by a unique form. On the other hand, a *maximally efficient* language would have a single, easy to learn and use form (perhaps the form, *ah*). The fact that language users need to be both effective *and* efficient requires natural languages to find a balance between these two opposing factors, as has been long discussed by functional linguists (Briscoe, 1998; Bybee, 1985, 2003; Givón, 1979; Goldberg, 1995; Grice, 1975; Haiman, 1985; Levinson, 1983; Paul, 1888; Slobin, 1977; von Humboldt, [1832] 1999).

The recognition that languages must be efficient and expressive, and that these pressures are mutually constraining, has gained new traction within the "noisy-channel" approach to language processing, which recognizes that speakers are attempting to express information as efficiently as possible, under imperfect or noisy conditions (Gibson et al., 2013; Jaeger and Levy, 2006). The noisy-channel approach has emphasized the dynamic nature of the balance between efficient and expressive communication. When a speaker is reasonably certain that an intended message will be successfully conveyed, the balance is tipped toward efficiency, with forms being reduced and distinctions being underspecified (Hopper and Traugott, 2003; Jaeger, 2010; Levy, 2008; Levy et al., 2009; Piantadosi et al., 2011, 2012). For example, when a verb appears with the construction it most commonly appears in, the verb form itself tends to be reduced (Gahl and Garnsey, 2004). Similarly, the complementizer *that* is more likely to be omitted when it is predictable in context (Wasow et al., 2011). On the other hand, under noisy or uncertain conditions, distinctions may be exaggerated, and language may be made less ambiguous in various ways (Bradlow and Bent, 2002; Buz et al., 2014; Gibson et al., 2013; Hall et al., 2013). Thus, efficiency and expressiveness balance each other and lead languages to vacillate between using shorter or fewer forms to express a given message on the one hand, and

adopting new, longer, or additional forms in order to ensure that messages are understood as intended, on the other.

The last idea, that speakers tend to obey the conventions of their language community, captures the fact that humans treat language as a normative enterprise. For example, people within a given community tend to believe that there is a "right" way to pronounce words, even if other communities are recognized to pronounce them differently. The Gershwins' famous lyric, "You like tomato and I like tomahto, . . . Let's call the whole thing off" epitomizes this idea. Humans are a rarity within the animal kingdom in using arbitrary communicative symbols that are shared within a community and distinct from those used in other communities (Tomasello, 2016). In fact, humans quite generally appreciate that there are culture-specific "right" and "wrong" ways to do a great many things, and we learn to obey these conventions in a way that other species do not (Boyd and Richerson, 1988; Horner and Whiten, 2005). Many normative conventions are, at least initially, self-conscious; for example, our knowledge about how to eat food politely, whether or how much to tip at restaurants, or whether it is polite to sneeze or burp in public. Other social norms may be obeyed without self-conscious awareness, including how close to stand to each other while speaking, or what sort of foods are appropriate for breakfast.

The importance of cultural norms for human behavior has enjoyed a long and rich appreciation within philosophy (e.g., Korsgaard and O'Neill, 1996). Our respect for normative patterns of behavior is what allows us to create complex cultural practices. For example, dollar bills would be meaningless were it not for the social agreement that imbues them with value. Driving would be terrifying if we couldn't rely on other drivers to (generally) obey standard driving practices. Again, work that has compared humans with other primates has emphasized that cultural norms may be uniquely human (Tomasello, 2009, 2016), particularly when they serve no clear purpose (e.g., Horner and Whiten, 2005; McGuigan et al., 2011).

Normativity is critical to the *explain-me-this* problem in that generations of learners obey restrictions that do not serve any clear communicative function: we respect the patterns that are evident in the input. For example, even though saying *she made it vanish* is somewhat less efficient than *she vanished it*, and even though the latter formulation is readily interpretable, native English speakers avoid the shortcut and normatively obey a shared preference for the periphrastic form. A mechanistic explanation supports our tendency to produce forms that our community deems acceptable, insofar as partially familiar formulations are easier to access than wholly novel formulations. That is, more conventional forms are more efficient to access from memory even if they are less efficient to produce (see section 4.5). But the reason native speakers judge the longer phrase (*explain this to me*) to be the "right" way to express the intended meaning

and the shorter phrase (?*explain me this*) to be "incorrect," and the reason that familiarity tends to be more important than ease of articulation, is because we desire to speak like others in our community—language is a social and normative enterprise.

Chapter 2 briefly reviews the nature of word meanings and asks how we learn to use words appropriately. We will then see in the following chapters that many of the lessons learned from an appreciation of word meaning extend naturally to our primary question: How do we learn to use basic clause types—ARGUMENT STRUCTURE CONSTRUCTIONS (ASCs)—in creative but constrained ways?

Chapter 2

Word Meanings

As described in chapter 1, the primary goal of this book is to provide an explanation of how native speakers of a language know that they can combine words in certain new ways but not others. I refer to this as the *explain-me-this* puzzle, since this phrase provides a good example of the phenomenon: the intended meaning of ?*explain me this* is perfectly clear, but it nevertheless does not sound natural to native speakers of English (hence it is preceded by "?"). Before addressing this issue, we first focus on the simpler case of how speakers learn to use individual words appropriately, because there are several key parallels between the two problems, insofar as both cases require learning how to use language in ways that are creative and yet constrained.

2.1 Meanings Are Rich, Structured, and Partially Abstracted

Many of us don't reflect on word meanings very often, and so we may fail to appreciate just how much intricate knowledge is required for us to use words appropriately. Quite specific knowledge is needed, for example, to distinguish the more than a dozen verbs that imply that a person is forced to leave a situation or place (see table 2.1). Each of these words is appropriate in a distinct range of contexts, even though they all designate the same abstract type of event represented in figure 2.1.

Several of the words in table 2.1 imply that the removed person has somehow transgressed; they differ in terms of which organization or place the person is removed from. *To banish* is to remove a person from civilized society; *to expel* (in one sense of the word) is to remove from a school; *to deport* is to remove

TABLE 2.1. Partial list of verbs that designate the removal of a person from a situation or place

Banish	*Dismiss*	*Extradite*
Blackball	*Evacuate*	*Fire*
Blacklist	*Exile*	*Kick out*
Cast out	*Expatriate*	*Lay off*
Deport	*Expel*	*Oust*

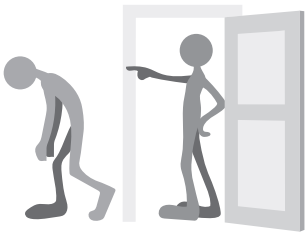

FIGURE 2.1. Concrete depiction of the general type of abstract event profiled by the verbs in table 2.1.

from a country; *to fire* is to remove a person from his job. *Lay off* differs from *fire* in that there's no stigma attached to the person who has to leave his job. A person who is *blacklisted* is barred from some type of work, often unfairly and for political reasons. People can be *ousted* only if they are political figures. This sort of contextual information or world knowledge is part of our knowledge of language, an observation that will be relevant as we approach the *explain-me-this* puzzle in the following chapters.

Different types of world knowledge are needed for different types of word meanings. Knowledge of quite specific motor sequences is required to distinguish the meanings of *walk*, *skip*, *hop*, *jog*, *power walk*, *run*, *race*, *sprint*, and *waltz*. While it would be difficult to try to explicitly define these terms, we can identify the actions at a glance. Moreover, we have contingent associations with many of these terms because of how they have been witnessed in the past. We know that *skipping* is most often done by children and is associated with being carefree, *power walking* is more commonly done in malls and by mothers or retirees, and *waltzing* is associated with formal events such as weddings.

In fact, once you start to think of the actual range of words that we know, it becomes clear that word meanings are not readily reducible to any finite list of recognizable features or attributes (Fodor et al., 1980) and word meanings are not easily disentangled from knowledge about real-world contexts in which these words tend to occur (Palmer, 1996; Willits et al., 2015). For example, if asked to try to define the word *bachelor*, many people will suggest "unmarried man" (Fillmore, 1975), a meaning that is often cited in philosophy classes because it seems to be definable in terms of simple features: unmarried and male. But this definition would seem to classify as bachelors the Pope, Tarzan, unmarried men with long-term live-in partners, and unmarried men on life-support. As Fillmore noted long ago, something clearly seems to be missing from this definition. He pointed out that the meaning of *bachelor* actually seems to be defined against a background of stereotypical world knowledge. This stereotypical knowledge takes for granted that boys grow up, date multiple people for a period of time, get married and stay married. Within this stereotype, there is clearly a period of time when an unmarried man is a bachelor. Fillmore suggested that each word meaning evokes a conventional SEMANTIC FRAME, where a frame is a structured abstraction or idealization over a set of coherent contexts (Fillmore, 1975, 1977, 1982, 1984; see also closely related ideas by Austin, 1962; Bartlett, 1932; Lakoff, 1987; Langacker, 1987). The word *bachelor* profiles an unmarried man, but it includes the relevant background frame of the stereotypical series

of events just mentioned. When the background frame does not apply, neither does the word.

Semantic frames capture rich aspects of our world knowledge as is required for adequate characterizations of word meanings. Experimental work has found support for the psychological relevance of contextually evoked semantic frames of knowledge. For example, Bar (2004) has demonstrated that people perceive the same visually filtered shape to unambiguously be a hairdryer, if placed in the context of a bathroom sink, but to be a drill if placed in the context of a workbench (see also Beck and Kastner, 2009; Biederman et al., 1982; Murphy and Wisniewski, 1989; Walther et al., 2009; Zettersten et al., 2018). When we interpret individual words, neural areas that relate those words to various actions or perceptions may also be reliably activated, suggesting that our sensory knowledge, whether directly experienced or imagined, is linked to word meanings. More specifically, words referring to actions performed by the leg, arm, or face (e.g., *kick*, *pick*, or *lick*) have been found to evoke activation at or near the areas in the primary motor cortex that are active when we move our feet, hands, or tongue (Hauk et al., 2004). Activation of our motor cortex is also evoked when certain words are used metaphorically—e.g., *to grasp the idea*, *to kick an idea around* (Cacciari et al. 2011; Desai et al. 2011; Pulvermüller et al., 2005). Texture-selective regions in the somatosensory cortex are likewise activated by words related to texture, whether used literally or metaphorically—e.g., *She had a rough day* (Lacey et al., 2012)—and cortical areas that are evoked by tasting food are activated when words related to taste are used in isolation or in conventional metaphors—e.g., *That was a bitter breakup* (Citron and Goldberg, 2014). In fact, different words reliably evoke activation that spans almost the entire cortex, suggesting that word meaning can involve a wide range of associations (Bergen, 2012; Huth et al., 2016). Thus, words evoke rich conceptual and perceptual information gleaned from the contexts in which the words have been witnessed.

2.2 Vast Implicit Memory

How do we *learn* the rich structured representations required for word meaning? One prominent philosopher found the question so challenging that he concluded that we must actually be born with all possible current, past, and future word meanings, so that our job as learners is merely to attach a label to an already existing concept (Fodor, 1975). But this idea leads to the uncomfortable stipulation that Abraham Lincoln was born with meanings (if not labels) for concepts such as *computer*, *fax*, and *drone*, and that your grandmother was born with knowledge of the meanings of *frack*, *twerk*, and *rap*. This seems unlikely.

It is clear that even a single initial encounter with a word can potentially leave a memory trace of its use. Why is this clear? Consider the counterfactual: if no

memory trace could exist after a single exposure, then the second time the word was encountered would be exactly the same as the first time. But then no memory trace of the word would be left upon this second encounter. This situation could be repeated ad infinitum without any memory trace of the meaning being retained. If this were the case, we would be utterly unable to learn any words. Therefore, it must be possible for an initial memory trace to exist in order for it to be strengthened upon subsequent exposure. Fortunately, we know that human brains have a vast capacity for implicit memory, even though memories may not readily be brought to consciousness (they are not always easy to *recall* or make *explicit*). In the domain of vision, for example, Brady and colleagues (2008) showed a group of participants 2500 images of distinct objects for three seconds each; remarkably, participants were able to recognize the pictures they had seen at well above chance rates, in that they were successfully able to select the picture they had seen from another picture of the same category (e.g., another dinner bell, star fish, or wooden desk). Participants were even able to distinguish the pictures they had seen from another picture of the *same object* positioned differently (e.g., a cabinet with doors closed or with one door ajar) (see also Standing, 1973). It seems that at least some of these fleeting memory traces become consolidated into long-term memory while we sleep (Marshall and Born, 2007; Stickgold, 2005).

An impressive capacity for memory for words has been found as well. Whenever we hear "the same" word form, it is pronounced slightly differently, as it may be spoken by a different speaker, or at a different speed, or in the context of different other words. While we readily generalize across these different tokens in order to appreciate that they involve the same word, it turns out we are faster to identify a repetition of a word if it is spoken by the same speaker, suggesting that aspects of that speaker's voice can be retained with the memory, at least for some period of time (Kleinschmidt and Jaeger, 2015; Palmeri et al., 1993). In fact, people find it easier to recall words from a list if they are tested in the same room in which they had originally read the list of words; this advantage can also be achieved if they are asked to imagine the original room just before recalling the list of words, as if the imagining is sufficient to reinstate the original context (Smith, 1979; see also Godden and Baddeley, 1975). In another experiment, people were shown a series of words (e.g., *spoon*, *table*, *house*) and pictures (of a fork, chair, shed) and subsequently queried, as they lay in a functional magnetic resonance imaging (fMRI) scanner to record their brain activity, whether or not they had seen a particular word or picture. In answer to a query about either *spoon* or *fork*, for example, the correct answer would be "yes." The queries were all made verbally; that is, the pictures seen initially were not displayed a second time. And yet the brain's *visual* cortex was found to be more active for those concepts encoded by pictures during the experiment (Vaidya et al., 2002). That

is, it appears that the word label used during the query (e.g., *fork*) evoked the visual memory of the picture (of a fork) if a visual image was witnessed in the context of the experiment. Since very familiar words were used in the study, the findings demonstrate that the contexts associated with a word are continuously updated in a context-sensitive way (see also Goldinger, 1998; Hintzman, 1988).

Another way to make clear that we retain rich contextual information about how words are used stems simply from the fact that our vocabulary is rich and highly nuanced as already alluded to. This in itself provides evidence that we *must* retain quite specific contextual information about how words are used as we learn the meanings of those words (see also Borovsky et al., 2010; Johns et al., 2016; Nelson and Shiffrin, 2013; Walker and Hay, 2011). For instance, imagine that a child hears the word *write* for the first time in a sentence like *She's writing it now.* Does the child record the speed and care that is used to write? Exactly what was written? Whether the writing was done on paper or on a wall? Whether or not the writer was famous? Although it may seem unlikely that all of these aspects of the nonlinguistic context would be recorded, we *have to be able* to record detailed aspects of contexts as we learn the meanings of words, or else we would not be able to learn, without explicit instruction, the correct meanings of words such as *scribble* (write quickly), *scrawl* (write carelessly), *sign* (write one's name in cursive), *endorse* (write one's name in cursive on a check), and *autograph* (write one's name when one is famous). Thus, our memory for words is not only vast but also linked to the contexts in which those words are experienced.

At the same time, it is reasonable to assume that aspects of contexts that are perceived to be more relevant or that are more uniquely predictive of a given word are more likely to be retained (or are weighted more heavily). That is, surprising or unusual aspects of contexts are more likely to be encoded in memory than aspects that are ever present. Aspects of contexts and words that are more highly correlated (have higher MUTUAL INFORMATION) are also more likely to be retained. Of particular importance to children are the speakers' perceived intentions, as the child attempts to interpret what speakers mean by what they say (Austin, 1962; Carpenter et al., 1998; Clark, 1996; Levinson, 1983; Tomasello, 2001; Tomasello and Barton, 1994). In short, the representation of each encounter prioritizes information and predictive cues that are perceived to be relevant.

Moreover, memory for any experience is necessarily partially abstract insofar as experiences are not represented completely veridically. We can describe the representations of events as involving LOSSY COMPRESSION, by which we mean simply that not *all* information is retained. For example, we might have a memory trace of witnessing a kumquat that is abstracted away from the color of the kitchen table upon which it sat, the tiny scratch in its surface, and the length of its stem. As Christiansen and Chater (2016) emphasize, the rapid time scale of

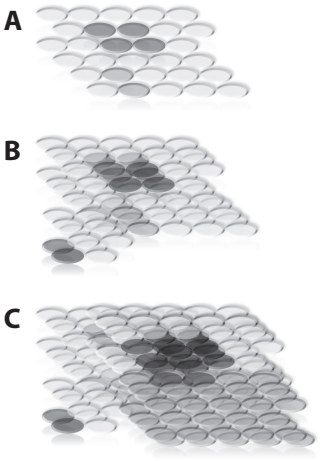

FIGURE 2.2. *A*, an abstract, structured, distributed representation of a single memory trace of a word witnessed in context; *B*, strengthened representation of the word as aspects are repeated over time (represented by darker nodes); *C*, representation of the word as it continues to be strengthened through experience.

language processing *requires* that our brains recode and compress incoming information.[1] Thus memory traces of experiences, no matter how vivid, are partially abstracted from our experience.

To summarize, an initial encounter with a word can form a lossy structured representation that prioritizes what the word designates and includes various contextual aspects of the encounter that are perceived to be informative or relevant to the use of the word, which may include quite detailed information about form, meaning, and context. Illustrated schematically in figure 2.2A is a single such representation, with more important aspects indicatd by darker ovals. As suggested by figure 2.2B, additional encounters with the same word typically overlap in some ways with the earlier representation, strengthening those shared aspects, while also potentially adding contextual information that is unique to that particular experience (Atkinson and Shiffrin, 1968; Light and Carter-Sobell, 1970). The result of multiple encounters of a word is a dynamic cluster of overlapping structured representations (figure 2.2C) situated within our HYPERDIMENSIONAL CONCEPTUAL SPACE. As indicated in the changes from figure 2.2A to 2.2C, the representation of a word becomes broader as aspects of each context are added. At the same time, those aspects of memory representations that overlap across multiple encounters of a word become strengthened over time, thus becoming more central to a word's meaning. In figure 2.2, overlapping aspects are indicated by the progressively darker parts of the representation. Figure 2.2 is also intended to convey that as memory traces accrue, their status as unique encounters gives way to an EMERGENT CLUSTER (or "cloud"), which constitutes what we think of as a single coherent word meaning (or *lemma*).[2]

1 Christiansen and Chater seem to imply that the "continual deluge of linguistic input" acts to "obliterate" previous linguistic material due to interference (2016, 1), but it is clear from conversations with both Morten Christiansen and Nick Chater that they fully recognize that memory traces of language experiences at different levels of generalization are retained in long-term memory (as long as they are consolidated, which often requires a night's sleep), as just emphasized.

2 Although the representations in figure 2.2 portray tight clusters in a small space, the clusters of neurons that are evoked by individual words typically span multiple brain areas.

The required representational space for language is hyper-dimensional and it must be part of our conceptual system because such an open-ended array of contextual factors is required for our representations of word meanings. That is, as described above, word meanings involve complex relational structure and do not correspond to lists of recognizable features (try to decompose the word *extradite* into a list of features and you'll appreciate this fact).

The formal domain (e.g., of sounds, word forms, word order) is quite a bit less variable than the open-ended domain of nonlinguistic context. Because of this, the formal similarities among representations play a special role in determining how our representations of experiences with words cluster together. Individual words are represented by a cluster of abstracted sequences of sounds and structured context-based semantic representations. The cluster emerges on the basis of similarities and parallels within representations, and from differences between other existing clusters. Thus, a "word" is in fact a cluster of partially overlapping structured representations within our hyper-dimensional conceptual space (see also Elman, 2009; Kapatsinski, 2018; Pierrehumbert, 2002).

At the same time, the conceptual space that is used for language does not cover our *entire* perceptual and conceptual knowledge or capacity: we have a myriad of nuanced experiences of smells, feelings, visual knowledge, and autobiographical memories that are not necessarily captured by any conventional words or constructions. In addition, as we will see, there exist dimensions that may be highly relevant to linguistic representations in a few languages. Therefore, the high-dimensional conceptual space used for language is a subset of our even more complex perceptual and conceptual system.

The representations in figure 2.2 are intended to capture the fact that each word's presentation is affected by how often it is encountered (its TOKEN FREQUENCY) and the range of contexts in which it has been encountered. And, indeed, it is well established that words that have been encountered more frequently are reliably faster and easier to access in a range of tasks (Broadbent, 1967; Jescheniak and Levelt, 1994; King and Kutas, 1995; Meyer and Schvaneveldt, 1971; Oldfield and Wingfield, 1965; Rayner and Duffy, 1986). We even have explicit intuitions about the relative frequencies of words: we recognize that *elephant* is more frequent than *pachyderm*, and that *dog* is more frequent than *elephant*. People are also able to judge quite accurately the relative frequencies of words encountered within experimental settings (e.g., Balota et al., 2001; Brysbaert and New, 2009; Hintzman, 1988).

The variety of contexts in which a word has been witnessed also appears to play a key role in how quickly and accurately that word is accessed (Adelman et al., 2006; McDonald and Shillcock, 2001). Words that have only occurred in a narrow range of contexts may be restricted to those contexts. For example,

no matter how many times we may have read Dr. Seuss's *Yertle the Turtle*, none of us consider *Yertle* to be a popular name. Instead, it immediately evokes the Dr. Seuss book. When a British friend says she is *chuffed*, I may understand its intended meaning in the context ("pleased"), but I am unlikely to use the word myself because I will associate the word with a different dialect of English.

This perspective predicts that knowledge of a word is not an all-or-nothing affair. The ability to correctly identify the meaning of a novel word in context (e.g., Carey and Bartlett, 1978; Woodward et al., 1994), does not entail the ability to successfully recall the meaning or even recognize the word after a delay. In one study, Horst and Samuelson (2008) demonstrated, for example, that two-year-old children were able to assign four new labels to novel objects when each new word was presented alongside two objects with familiar names (e.g., a pair of glasses, a toy dog, a "cheem"). But the children showed no evidence of retaining the word-object mappings after a five-minute delay when the new object (e.g., the "cheem") was presented along with two other novel objects. Therefore, although we retain a great deal of implicit information upon initial encounter, our ability to actively make use of the word-level abstraction over contexts develops through time and experience with the word (see also Fernald et al., 1998).

2.3 Clusters of Conventional, Related Senses

Most of the time, new contexts are relevantly similar to familiar contexts and so we can simply use our learned words to convey familiar meanings. But sometimes we encounter importantly new contexts, in part because the world itself is constantly changing, and in part because our relation to the world is changing. When faced with especially new contexts, we may need to use old words in creative ways. For example, the terms *file*, *folder*, *paper*, and *trash* all took on new meanings when they were applied to computer interfaces; they were immediately interpretable because they relate in clear ways to their original senses. A few other words that have acquired new conventional senses in addition to their original senses are provided in table 2.2. The employment of an old word for a new purpose is efficient for the speaker, who does not need to coin a new term and risk communication failure; it is also effective for the listener, who is able to use knowledge of an existing sense of a word as scaffolding to a new sense. The familiar meaning provides an indication of what is intended. Initially, extended senses are created on the fly, but if a new use of a familiar word catches on in a wider community, the new sense becomes part of the conventional meaning of the word. This gives rise to conventional POLYSEMY: a network of related senses for individual words (Brocher et al., 2017; Copestake and Briscoe, 1995; Floyd and Goldberg, forthcoming; Geeraerts, 1993; Klein and Murphy, 2001; Lakoff, 1987; Tuggy, 1993).

TABLE 2.2. Familiar words that have acquired new senses relatively recently

Word	Prototypical original sense (still in use)	Newer sense
Bandwidth	The maximum amount of information that can be transmitted along a channel	The mental capacity of a person, as in, *I don't have the bandwidth to set up a carpool*
Lit	Verb: past tense of *to light*	Adjective: drunk or exciting (*I was lit*; *The party was lit*)
Friend	A close intimate that isn't family	Anyone given explicit access to a Facebook feed
Optics	A field of study concerning vision	How a public event is perceived
Spam	Brand name of mass-produced inexpensive and unappealing mystery meat	Mass-produced unsolicited and unappealing e-mail
Literally	True without exaggeration or metaphorical interpretation	Emphasis without being true, as in *We were literally killing ourselves laughing*[a]

[a] From the *Oxford English Dictionary*, 3rd edition.

For example, a prototypical sense of the verb *to fire* involves a very quick, violent triggering of a bullet from a gun that can lead to injury or death and is often performed out of anger. When someone is *fired* from a job, the quick and brutal implications of the prototypical sense of firing a gun are retained, but other aspects are not. We can also *fire off* a letter, which shares the sense of a quick action, performed with a sense of urgency or out of anger. When neurons *fire* we understand there to be a quick and directed event, but of course no volition, addressee, anger, or harm. If we were to try to distill the meaning of the verb *to fire* into what all senses have in common, we might come up with a meaning such as "any quick, directed action." But this meaning would be too general, since it would seem to imply that chopping a tree or breaking an egg into a bowl could be instances of firing, and it would also fail to capture the implication of most of the senses (but not the firing of neurons) that harm is intended. Rather, multiple senses often cluster around a rich, prototypical (often stereotypical) semantic frame, with extensions based on some but not all of the attributes of this frame (Lakoff, 1987). The senses just mentioned are described

TABLE 2.3. Related senses of the verb *to fire*

Original prototype on which other conventional senses arose historically: *to fire* a gun	The quick action of pulling a trigger on a gun in a directed way, with the intention of causing a bullet to strike something; can be and often is repeated in rapid succession; often done out of anger and with the intention of causing physical harm or death
To fire someone	To lay off someone from employment, typically resulting in financial and emotional harm to the employee; done quickly (without warning); sometimes done out of anger
To fire questions or insults at someone	To communicate comments in a quick and harsh way
Neurons *fire*	The action potential of a cell sends an electrical signal down the axon (quick, directed)
To fire off a letter	To quickly direct a communicative act toward someone, typically out of anger

in table 2.3 and represented in figure 2.3.[3] The observation that words are commonly associated with a "radial category" of senses radiating out from a rich prototypical semantic frame is due to Lakoff (1987), who argues that clusters of such senses exist for most nouns and verbs, except those that are used infrequently or in restricted technical contexts (Dautriche, 2015; Durkin and Manning, 1989; François, 2008; Heylen et al., 2015; Tuggy, 1993). To summarize, it is clear that words convey rich frame-semantic meanings, typically associated with a network of related meanings because new senses arise from the need for words to apply to new contexts.[4]

3 The prototypical meaning of the verb *fire* is also related to the noun *fire* (not pictured), since the firing of weapons commonly involved a flame. The noun *fire* is directly related to additional verbal senses including *to fire clay* and *to fire up a conversation*.

4 Over time, an original meaning may die out, leaving only the newer sense(s). For example, *divest* used to mean *undress* or *deprive others of rights or possessions*. These days, of course, it is restricted to the selling off of investments of a certain type. The word *clue* used to refer to a ball of yarn; since we can talk about "following a thread" to figure something out, the new and old senses are likely related but the original meaning of *clue* has been lost. More

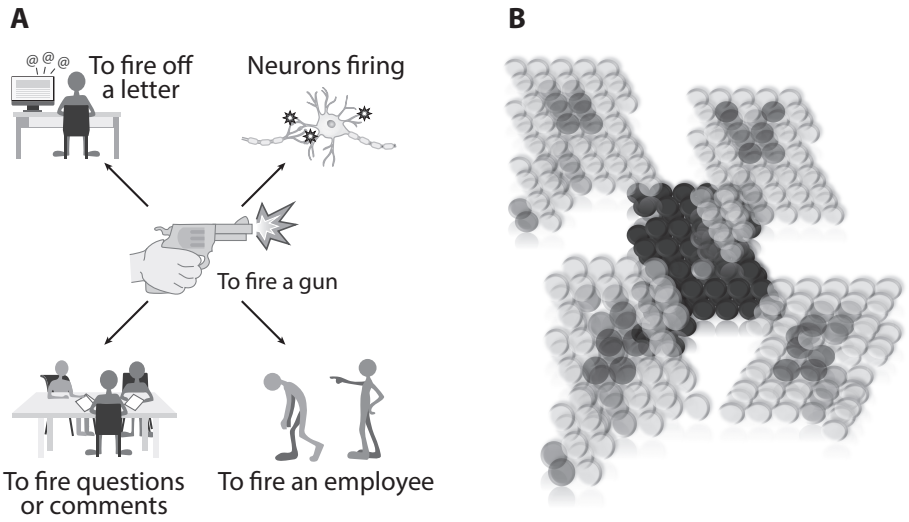

FIGURE 2.3. *A*, symbolic, and *B*, abstract, structured, distributed representation of related senses of the verb *to fire*.

2.4 Creativity

Assigning meaning to words is recognized to be a difficult and complex task (Gleitman, 1990; Quine, 1960). This has led to the idea that meanings that are initially associated with a word are routinely incorrect, and in need of correction or elimination. One popular idea is that learners form a tentative hypothesis about what a new word means, and then either confirm or reject that meaning when they encounter the word in a new context. If the new context is inconsistent with the hypothesized sense, this thinking went, the hypothesized sense would be dismissed, with learners essentially deleting the initially proposed association between the word and the original sense (Aravind et al., 2018; Berwick and Weinberg, 1986; Siskind, 1996; Trueswell et al., 2013; Woodard et al., 2016).

But armed with a fuller appreciation of word meanings, we can see that this "Propose but Verify" model, as it was called, cannot possibly work, because words typically have multiple distinct senses. Encountering a second sense of a word cannot eliminate the first sense. For instance, hearing *fire* used in the context of employment cannot lead the learner to eliminate the sense of *fire* associated with guns. It cannot even lead the learner to delete particular aspects of the hypothesized senses, because multiple senses of a word do not necessarily share any particular attribute among them. For example, consider the word *breakfast* (Fillmore, 1976). If a child initially encounters the word *breakfast* in the

commonly, old and new senses coexist as we've seen in the case of the multiple senses of *fire*, and all of the examples in table 2.2.

context of being told to *eat your breakfast*, the child may (correctly) hypothesize that *breakfast* refers to "food, like this cereal, that constitutes the first meal of the day." But what if a child later hears that she can order breakfast in the evening at a diner? Should she jettison her original hypothesis? Clearly not. Should she delete the restriction on the time of day and retain "food such as cereal"? The latter interpretation would be sufficient for these first two contexts, but what is the child to think when she subsequently hears that someone had pizza for breakfast? Adding this third context would require eliminating the hypothesis that *breakfast* refers to either "food such as cereal" *or* "food that constitutes the first meal of the day." But then the child is left with no meaning for *breakfast* at all!

The Propose but Verify model has since been updated by a "Pursuit" model, to allow for hypothesized meanings to be incrementally suppressed, rather than automatically eliminated, whenever a hypothesized meaning is incompatible with the current context (Stevens et al., 2017). Yet, oddly enough, this model stipulates that only one potential sense for a word is considered:

> If the most probable hypothesis fails to be confirmed, the model does not test out the second most probable hypothesis, but rather chooses a meaning from the current context instead. Furthermore, the selection of the new meaning also follows a minimalist strategy: If there are multiple meanings available, the model does not favor meanings it has seen before but chooses completely randomly. (Stevens et al., 2017, 12)

For this reason, the Pursuit model, like the Propose but Verify model before it, fails to allow words to be systematically associated with multiple related meanings (see Floyd and Goldberg, forthcoming).

Instead of deleting or suppressing a hypothesis that *breakfast* means "stuff like this cereal that is eaten in the morning," a child would be better off retaining this sense and *adding* additional senses, including "food, like this cereal, regardless of when it is eaten" and "food, regardless of what it is, that constitutes the first meal of the day." The first context instantiates the most common or prototypical sense of the word, after all, and the latter two senses are conventional extensions of that sense. The suggestion then is that memory traces of word forms and contexts are only added; an association between a word and meaning cannot be eliminated by additional encounters with the same word.

For the most part, past the age of a year and a half, learners are remarkably accurate at assigning the correct meanings to novel words, as they make use of a variety of cues to determine intended meanings (Akhtar et al., 1996; Baldwin and Tomasello, 1998; Childers and Tomasello, 2002; Imai et al., 1994; Landau et al., 1988; Markman, 1989; Mervis et al., 1994; Soja et al., 1991). In assigning

an interpretation, even young children dynamically take into account not only their own knowledge of the discourse context, but also what they take to be the speaker's current state of knowledge. For example, Tomasello and Haberl (2003) found that 12- and 18-month-old infants, who were equally familiar with and interested in three toys, tended to offer the one toy that was new to an adult, when she asked, "Oh, wow! That's so cool! Can you give it to me?" (see also Gweon et al., 2014). That is, even very young children recognized that the adult was likely interested in the toy that was novel to that adult. Still, of course, learners can and do make errors of varying types. Can the idea that contexts are only added actually work?

2.5 Competition Constrains Word Meanings

How can a child ever *unlearn* a sense that is incorrectly associated with a word? For example, what if she incorrectly assumes that *breakfast* means "cereal"? Witnessing *breakfast* in additional contexts of cereal would not dissuade her of this sense, and we have just argued that witnessing *breakfast* in contexts *without* cereal should also not dissuade her of this sense. How, then, can the child come to dissociate the meaning of *breakfast* from the more specific meaning "cereal"?

In fact, the word *cereal* must come to the rescue. In contexts in which cereal but not breakfast is intended, the child will reliably hear the word *cereal*. The new association between *cereal* and "cereal" will eventually become stronger than the previous association between *breakfast* and "cereal." That is, each word fills its own semantic and distributional niche (Aronoff and Lindsay, 2016; Clark, 1987; Gauger, 1973). Words are not interchangeable and true synonyms are vanishingly rare. Each word is distinguishable from other words because each is associated with a rich network of contexts or semantic frames, as already discussed. Near synonyms reliably have different contextual restrictions or background frames; they may differ, for example, in terms of formality (*dog* vs. *pooch*), perspective (*ceiling* vs. *roof*), or attitude (*skinny* vs. *slim*).

Since we have seen that our memory for words includes aspects of their contexts of use, it is reasonable to assume that intending the same type of message in a relevantly similar context will serve to activate words used previously in those types of contexts to express that type of message. Whenever multiple words are activated and cannot be combined, they compete.

Thus, if children initially incorrectly assign a meaning to a word, they will eventually learn a different term that is more appropriate for the intended meaning. In this way, word forms are in competition with one another for meaning: *cereal* will eventually STATISTICALLY PREEMPT the incorrect association between *breakfast* and "cereal." If there is no more appropriate term, then there is no need for children to modify their sense of the word, and there is no evidence

that they do. Instead, appropriate interpretations of a word will be strengthened over time as the word is witnessed in additional contexts.

The fact that words influence the potential meanings of other words is clear if we consider examples from other languages. For example, in German, the word *Blase* means "bubble" and is used to refer to blisters as well as soap bubbles. But among English speakers, only children or second-language users are likely to refer to a blister as a *bubble* because the word *blister* exists and is a more appropriate alternative. Ancient Japanese used a single word, *ao*, to describe either blue or green color, but English speakers cannot call the sky *green* or the grass *blue* because the other term is better suited to convey the intended meaning. Conversely, we rather indiscriminately speak of *knowing* how and *knowing* something, whereas French distinguishes the two types of knowledge with distinct verbs, *savoir* and *connaître.*

The hypothetical *breakfast* for "cereal" example discussed above is a case of a child initially *under*generalizing the meaning of a word; but what if she initially *over*generalizes the word's meaning? In the early stages of word learning, young children do overuse words in their small vocabularies in ways that are not conventional in the language (Bloom, 1975; Clark, 1973). For example, children may use the word *ball* to refer to a button or the moon. They may use the word *dog* to label all animals, and they may say *hi* when they should say *bye*. These overextensions are most common when children's vocabularies consist of only 50–150 words (Gershkoff-Stowe, 2001), an indication that they simply do not have a better word for the meaning that they wish to express at that particular moment. In an important study, Kuczaj (1982) found that when asked to "show me the dog" or "show me the ball," the same children who overgeneralized these words in their own utterances reliably chose a picture of a dog over pictures of other animals and chose a picture of a ball over pictures of other round things (see also Gelman and Naigles, 1995; Huttenlocker, 1974). That is, children seem to know that, at least, the best examples of the category *dog* are dogs, not other animals, and the best examples of *ball* are actual balls, not other round things. This is expected, given that that memory traces of encounters with words are retained: children may initially use words in an overly broad way, but they recognize that situations which better match their representations are more appropriate. Thus, children produce overgeneralizations when no other word has been learned or is sufficiently accessible at the moment of speaking (see also Harmon and Kapatsinski, 2017).

Overgeneralizations, just like undergeneralizations, are eventually avoided as new words enter the child's vocabulary. The child stops calling the moon *ball*, not because she witnesses *ball* in additional contexts, but because she learns and accesses the more appropriate word, *moon*: *moon* statistically preempts the use of *ball* to refer to the moon. What we learn are correlations between

word forms and a range of rich representations that capture senses in contexts. The suggestion then is that different word forms are in competition with one another for a particular construal: if the word *moon* is used to refer to the moon in certain types of contexts, then *ball* is not used in to express that meaning in those types of contexts.

In the following section, we review the evidence in support of this idea. The notion of competition via statistical preemption will also play an important role in our solution to the *explain-me-this* puzzle.

2.6 Learning and Fluency Reduce Overgeneralizations

Distinct words intended to express the same meaning-in-context compete to be used. There are almost always contexts in which one word is preferred over the other. For example, we may believe that *quick* and *fast* are interchangeable, but in fact it is much more natural to talk about a *quick shower* and a *fast car* than a *fast shower* or a *quick car*. This is because *quick* tends to be used when there is a clear end point, whereas *fast* readily refers to speed without invoking an end point. Since, as discussed above, the relevant contexts that become associated with particular words are quite rich, distinct labels are associated with distinguishable CONSTRUALS or MESSAGES-IN-CONTEXT, even when they are in certain ways quite close in meaning (e.g., compare *jock* vs. *athlete*; *thrifty* vs. *stingy*; *banish* vs. *deport*) (Langacker, 1987).

There is ample evidence for competition between word *forms* as words are being learned. By about one and a half years of age, if children are asked to find the *moop* among a group of objects, all but one of which already have a familiar label, toddlers will assume *moop* names the unfamiliar object. This tendency has been described as a bias to assume that a new word refers to a new object ("mutual exclusivity") (Markman, 1989; Markman and Hutchinson, 1984; Markman and Wachtel, 1988; Xu, 2002). But the same object in the world can often be labeled in multiple ways ([tele]*phone*, *cell*, *device*, *thingy*). The bias is more accurately described as resistance to using two distinct words to designate the same *construal* or *message-in-context*.

The idea that distinct words compete to describe the same message-in-context is sometimes confused with a quite different claim; namely, that we have a bias against one label being associated with more than one message-in-context. This latter claim does not appear to be valid, at least not when the messages-in-context are semantically related to one another (recall, again, the multiple distinct senses of the one word form *fire*) (see Floyd and Goldberg, forthcoming).

In the case of word learning, competition between word forms explains why we generally assume that a new word refers to a new distinguishable concept.

Various models have been proposed that take advantage of competition between word choices in both production (Bates and MacWhinney, 1987; Horst and Samuelson, 2008; Rayner and Springer, 1986; Yurovsky et al., 2013) and comprehension (Gaskell and Dumay, 2003; McClelland and Elman, 1986). The idea that words within a given dialect compete for distributional niches has a clear analogy in biology. Two species that share the same ecological niches cannot coexist in a long-term equilibrium; one or the other will gain a slight advantage, and this advantage will snowball over time, ultimately driving the other species either to extinction or to a newly distinctive ecological niche (Darwin, [1859] 1993; Grant and Grant, 2002). Darwin in fact long ago drew the analogy to language, noting that two words cannot remain in a long-term equilibrium if they are both associated with the same meaning. Aronoff and Lindsay (2016) suggest that the winner-takes-all competitive phenomenon extends beyond biology to all complex systems. They appeal to Gause's general law of competitive exclusion (Gause, 1934) as an explanation for the fact that languages strongly disprefer true synonyms. But words are not species and meanings are not food. So it is fair to ask *why* don't words exist that are in free variation, usable in identical contexts?

There is a functional advantage to assigning each label a distinct meaning-in-context, even if the meanings differ only in terms of register (*buy* vs. *purchase*), dialect (*pop* vs. *soda*), connotation (*stingy* vs. *thrifty*), background frame (*land* vs. *ground*), or distribution (*sofa* bed vs. therapist's *couch*). The fact that distinctions exist allows speakers to more quickly access the best match for their intended message-in-context when they speak. If two words were truly interchangeable, speakers would be forced to make a totally random decision each time either word was used. This would violate the efficiency aspect of the CENCE ME approach without contributing to expressiveness, since unbiased decisions take longer to make (Ratcliff et al., 2004) and yet contribute no additional information.

The idea that each word has its own distributional niche has been thought to be based on a pragmatic inference, in that the following reasoning is attributed to the learner: "if the speaker had intended to refer to the 'spoon' she would have used the word *spoon*; therefore, this word must not mean 'spoon'" (Clark, 1987; Diesendruck and Markson, 2001; Goldberg, 1995). While this type of inference may be made by older children and adults, there is no need to posit a high-level inference, or even one that is specific to humans, to account for the effect of competition between word forms. It turns out that well-trained *dogs* have been shown to prefer that a novel word be assigned a novel meaning; in particular, a border collie, Rico, was found to select the one new object from among several familiar objects that already had familiar labels, when asked to bring the *sirikid* (a novel term) (Kaminski et al., 2004). This effect can be attributed to classic interference: once an A→B association has been learned, it is more difficult to

learn a C→B association (Ellis, 2006). That is, once a word form (A) becomes associated with a particular meaning (B), it is more difficult to interpret an unrelated word (C) as having the same meaning. The ubiquity of polysemy predicts that learning an A→B association, however, does *not* make it more difficult to learn an A→B′ association, where B and B′ are distinct meanings that are recognizably related to one another.

2.7 Summary

There is much more to say about word meanings and how they are learned and used. The general topic deserves several monograph-length works of its own (see, e.g., Bloom, 2000; Bowerman and Levinson, 2001; Clark, 1995; Fellbaum, 1998; Hart and Risley, 1999; Lakoff, 1987; Murphy, 2002; Pustejovsky, 2012; Tomasello, 2003). However, our primary focus in this book is not on word meaning or word learning, but on the fact that learners are able to avoid using certain phrasal formulations even though those formulations are easily interpretable and are not syntactically ill formed in any obvious way. That is, our quarry is the *explain-me-this* puzzle. The key take-away lessons from the present chapter are the following:

- Words evoke semantically rich, structured meanings, partially abstracted from contexts of use.
- Memory for how words are used is vast.
- We regularly employ old words for new uses, so that common words come to evoke a cluster of conventional, related senses.
- New representations are added, strengthening the overlap in our hyperdimensional conceptual space with previous experiences of the same word.
- Word meanings are constrained by competition from other words.
- Speakers avoid overgeneralizations by learning and gaining fluency with more appropriate labels for the intended meanings.

In the following chapters, we will see that each of these points applies to more abstract pairings of form and function, including ARGUMENT STRUCTURE CONSTRUCTIONS, introduced in the following chapter. We address the *explain-me-this* puzzle head on in chapters 4–6.