The Psych 101 Series

PSYCHOLINGUISTICS 101

H. Wind Cowles

SPRINGER PUBLISHING COMPANY

Psycholinguistics 101

H. Wind Cowles, PhD, is an Assistant Professor in Linguistics at the University of Florida. She received her PhD in Cognitive Science and Linguistics in 2003 from UC San Diego and was a research fellow in Experimental Psychology at the University of Sussex with Professor Alan Garnham. Her main research focus is on the role that discourse and information structure play in the written comprehension and spoken production of sentences. Her research applies multiple research techniques, including eye-tracking and event-related brain potentials, to questions about the ways in which people make reference and build structure in language. She is committed to both research and teaching and received a Teacher of the Year Award from UF's College of Liberal Arts and Sciences in 2009. Her training and experience in psychology, linguistics, cognitive science, and neuroscience make her part of a new wave of researchers approaching psycholinguistic research from multiple perspectives and methods.

Psycholinguistics 101

H. Wind Cowles, PhD

SPRINGER PUBLISHING COMPANY
NEW YORK

Springer Publishing Company, LLC
11 West 42nd Street
New York, NY 10036
www.springerpub.com

Acquisitions Editor: Nancy S. Hale
Senior Editor: Rose Mary Piscitelli
Cover design: Mimi Flow
Project Manager: Ashita Shah
Composition: Newgen Imaging

ISBN: 978-0-8261-1561-4
E-book ISBN: 978-0-8261-1562-1

10 11 12 13 / 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Cowles, H. Wind.
Psycholinguistics 101 / by H. Wind Cowles.
p. cm.
Includes bibliographical references.
ISBN 978-0-8261-1561-4
1. Psycholinguistics. I. Title.
P37.C69 2010
495.1′5—dc22 2010026859

Printed in the United States of America by Hamilton Printing

Contents

Acknowledgments

This book is a long time in coming and has spanned not one or two but three editors. So, I'd first like to thank Phil Laughlin, Sheri Sussman, and Nancy Hale for their help, encouragement, and infinite patience with a novice book writer.

This book benefited from the help of several of my students, and I would like to particularly thank Aazam Feizmohammadpour and Sangyeon Park for their help in gathering citations and tracking down research. I would also like to thank the members of the Language and Cognition Lab and the Language and Brain Interest Group at the University of Florida for interesting and insightful discussion on many of the topics in this book.

I was also fortunate to get feedback on some of the chapters from some experts in the field—thanks to Edith Kaan, Jamie Reilly, and Robin Thompson for their insightful comments, helpful suggestions, and the occasional gentle correction. Any mistakes remain my own, of course.

Finally, I'd like to thank Esme Cowles for providing his "nonexpert" expert advice and comments on the book, and his all-around support and encouragement.

1

Introduction: What Is Psycholinguistics?

DO YOU KNOW HOW YOU ARE ABLE TO READ THIS SENTENCE?

Language serves a central role in our daily lives, be it talking about complex thoughts and ideas, ordering lunch, reading a magazine, or persuading someone to do a favor. It's hard to imagine going through an entire day without using it—without reading, producing, or understanding a single word. Picture for moment what it would be like if you needed to get a ride from a friend to the airport next Saturday but couldn't use language in any form to communicate—how would you ask? How would you know that your friend understood and would give you a ride at the right time (or not)? For that matter, how would you have booked your flight in the first place?

Language is everywhere in human society, across every culture. For most people, it is an effortless ability that we acquire before we learn how to dress ourselves. Some people feel that they even think using language, that an important part of their thoughts involves a kind of internal monologue. Yet, language is not a simple thing—it is actually quite complex and we need to be able to use it quickly. In fact, its timing is so crucial that even minor delays during the processing of language are thought to have profound consequences. So, by understanding how language works and how we are able to use it, we can understand a vital yet complex part of our daily lives, offer ideas for how to help people with language disorders, and perhaps even shed light on how we think.

ON BEING A "PSYCHO" LINGUIST

What is a psycholinguist? As the name suggests, it is someone who studies phenomena in the intersection of linguistics and psychology (or a deranged linguist, of course!). But this does not really answer the question, it just pushes back. So, first—what is linguistics? It's the scientific study of language. Psychology? The scientific study of human behavior and cognition (i.e., how we think). The union of these two fields is principally concerned with the processing and knowledge representations that underlie the ability to use language, and how they relate to other aspects of human cognition. In short, psycholinguistics asks the question: How it is that people are able, moment-by-moment, to produce and understand language? And, by extension, how do children come to have this ability? How and why it is sometimes impaired after brain damage?

The term "psycholinguistics" is a little bit misleading in that it minimizes the contribution of about a half-dozen other fields of scientific research. To truly get to the bottom of how language works, we need expertise not only from linguistics and psychology but also from other scientific fields such as neuroscience and computer science. The whole endeavor of psycholinguistics often finds a home in the broader research field of cognitive science—an interdisciplinary field that addresses the difficult question of how animals, people, and even computers think.

Why the need for at least four research fields to understand something as seemingly trivial as language—something that most people have pretty much mastered before they turn four? As I mentioned earlier, language is more complicated than it might first appear, and so it is the very ease of use that makes it so interesting. Let's start with an analogy—walking around is also pretty easy for most people and it is something that we usually learn to do even before we learn how to talk. But can you explain exactly how it works—how each muscle, bone, and tendon allows us to defy gravity at every step? It's easy to do but hard for a layman to explain how it works—instead it takes an army of biologists, physiologists, and others to explain this "easy" skill. Language is very much the same—its function is easy to master (when you're young), but hard to explain. And, like walking, we have a lot to gain by understanding it.

The centrality of language in our daily lives means that any disruption to our ability to use it may be keenly felt—the worse the disruption, the more devastating the impact. This disruption can come as part of aging, it can come from developmental problems, and it can come from damage to the brain (or other body systems that impact our ability to perceive or produce language, such as damage to our hearing). In all of these cases, the disruption may be relatively small—perhaps difficulty in finding the right word or understanding very complex or unusual language. But, it can also be quite severe—from a complete loss of nearly all language to a particular difficulty with production, or understanding even fairly simple sentences. Of course, research on how this breakdown occurs can help us better understand how to help restore language function. But even research on how language processing works in a fully functioning system is very important because it can provide us with a model of language processing that can then be used to develop more effective therapies for people with disruptions to their language ability.

CHAPTER 1

THE NEED FOR SPEED

One of the key challenges for processing language is that it comes at people fast. While speaking rates differ somewhat among speakers and languages, on average people have to deal with around 2 to 3 words per second—that's roughly 5 to 6 syllables or 25 to 30 individual sounds per second. Perhaps this is not striking to you—clearly, it's easy because everyone can do it, so how is this a challenge? To convince yourself, go right now to a computer and visit a Web site like the BBC World Service that plays audio broadcasts in languages from all over the world. Pick a language that is unfamiliar to you, and listen to some of the broadcast in it. Of course, it's not going to have meaning for you, but let's set that aside for the moment. Focus instead on the sound of the language. Can you make out the individual sounds? Probably you can catch some of them, but could you write down each sound that you hear? Can you tell where all the word breaks are? Probably not.

So, language is easy when you are a native or fluent speaker of it. But, if you don't know the language, it's hard to even catch all the sounds. Clearly, there is something critical to being a speaker of a language that allows very rapid information processing.

Let's turn back to a language you do know—English—and think about all that's involved in successfully understanding just a single sentence. We can use the following sentence in 1 as an example:

1. You called Frank about the computer yesterday?

Suppose you hear this question (rather than reading it)—let's start at the most basic steps and work up. First, you need to be able to make out the sounds that you hear, and break them up into words. Next, you need to be able to assign meaning to those words, for example, you need to know who said the sentence and to whom they said it in order to know who the "you" and "Frank" are, you need to know what a computer is, and so on. Next, you need to know who did what—in this case who did the calling. There is also information about when this event happened, and some extra information about the topic of the event (it involved a computer in some way). Next, the intonation of the sentence (marked in writing with a question mark) tells us that this is a question, not a statement, and this (finally) leads to the fact that to fully understand the sentence we must link it to our wider understanding of the world and the intentions of the speaker. The speaker may be looking for confirmation about something he/she can't quite remember, can we provide it? This also means knowing what it was about the computer that would prompt a call to someone named Frank.

In the rest of this book, we will examine many of the research findings from psycholinguistics that tell us something about how these steps work, and more. First, however, in this chapter we will start with a brief history of the field to situate the current state of things and then we will discuss some of the overarching issues that frame much of the research. Finally, we will cover a brief overview of the rest of the chapters in this book.

A BRIEF HISTORY OF PSYCHOLINGUISTICS

Interest in the mind and language both date back for millennia, with a documented history of language study going back 2,500 years and spread across many cultures (including India, China, Mesopotamia, and Greece). Documented interest in the mind and knowledge—the foundations of what we consider to be psychology—also dates back, in one form or another, at least as early as language study, and perhaps earlier. However, modern versions of both linguistics and psychology are much more recent, with modern psychology tracing back to Wilhelm Wundt's lab in Leipzig in the late 1800s (we'll be discussing Wundt's work shortly) and modern linguistics tracing back to roughly the same time. Both fields have undergone some revolutions in even that relatively short time, with both fields experiencing some major shifts in the mid-1900s that are still felt today.

From the beginning of psychology, there has been an interest in language. Wilhelm Wundt, for example, published a book on language (*die Sprache*) in 1900. This book, with 1,367 pages by its 1913 edition, covered a number of topics that are still very much relevant in current psycholinguistics, including child language acquisition, sign language, language perception, and grammatical structure. The interest between the domains of language and psychology was mutual and, as Blumenthal (1987) discussed, many linguists of the day were also interested in Wundt's work and attended his lectures at the University of Leipzig, including several influential language researchers such as Leonard Bloomfield, Ferdinand de Saussure, and Edward Boas. For example, Bloomfield's approach to analyzing the structure of language is important to almost all modern theories of grammar. De Saussure made a critical distinction that is still part of how language researchers think about language: *langue*, which is the knowledge of a language system that exists collectively among speakers of a language, and *parole*, which is the use of that system (see Seuren, 1998, for an excellent history of linguistics).

While Wundt was an experimentalist (in fact, famously so—he is considered by many to be the father of modern experimental psychology), he also acknowledged the importance of internal mental states and viewed language as importantly reflecting mental representations. He viewed the sentence as a key unit of language, and sought to show how universal characteristics of human information processing, like attention and memory, would influence its production and comprehension. This is really not so different from what modern psycholinguistic researchers are doing, though the route from Wundt's experiments to the present day is not entirely direct.

Wundt was not the only researcher at that time interested in language, of course, and in fact there was some conflict between Wundt's approach and others, including approaches that disagreed with the idea of a unified mental representation and other types of "introspective" approaches. In fact, at the turn of the century there were quite a number of competing perspectives on psychology, which led one linguist of the day, Bernard Delburck, to suggest that linguists might do better to part ways with psychology. This is largely what happened, and for the next 30 to 40 years linguistics focused instead on the formal aspects of language—sound systems, grammatical structures, word formation rules—without much reference to the mental processing needed for their actual use. This approach still forms a core aspect of linguistic study today.

Although Wundt's work clearly foreshadows modern views and topics on psycholinguistics, his influence declined following the first world war. Bloomfield, once a proponent of Wundt's approach, had turned to behaviorism instead by 1933 when he published one of his major contributions to linguistics, a book simply called *Language*. In psychology, behaviorism was a movement in which the study of mental states was more or less rejected, and the idea that one could account for human behavior in terms of mental states or representation was discounted. In linguistic terms, this meant a stronger focus on descriptive accounts of language rather than studying language as a window onto human mind. Perhaps the most famous attempt to account for language processing in a behaviorist tradition comes from B.F. Skinner's 1957 *Verbal Behavior* in which language is not a complex mental construct with rules and representations, but instead is reduced to, well, verbal *behavior*. As such, it can be explained, according to Skinner, in terms of the same conditioning theory that applied to other behaviors, such as classical and operant conditioning, in which links between stimuli and outcomes are formed and shaped by experience. For example, a child saying "I want milk" may result in the child receiving a glass of milk, and this reinforces (or conditions) the use of this verbal behavior.

The trouble was that there are a number of aspects of language that cannot be explained by classical and operant conditioning. In a famous critique of Skinner's book, Noam Chomsky (1957) successfully argued against verbal behaviorism with several key points. Crucially, as we shall see in chapter 2, language is recursive and can produce an infinite number of sentences from a finite set of systematic rules and representations. The complexities of language that these rules create are difficult (if not impossible) to account for in simple stimulus–response terms. As part of his arguments, Chomsky reintroduced the idea of mental representations back to the study of language. He also drew an important distinction between the knowledge that one has about a language, called "competence" and the use of the language, "performance" (similar to the distinction of *langue* and *parole* drawn by Ferdinand de Saussure roughly 60 years earlier). Chomsky's influence on modern linguistics and psycholinguistics is profound, and his focus on competence (as opposed to performance) drew linguistics heavily in this direction. On the other hand, psychology continued to be quite interested in the concept of language performance. Nonetheless, several of Chomsky's proposals about the nature of syntactic structure, and in particular his work on transformation grammar, prompted experimentation by psychologists in the 1960s to see whether the linguistic processes proposed were psychological processes. For example, one could reasonably ask whether structures that were proposed to be more complex linguistically would cause longer processing times. The results were mixed—research showed that there was an important relationship between linguistic structure and psychological processing, but didn't support the particular relationship proposed by transformational grammar, despite initial successes (e.g., Miller & McKean, 1964; Slobin, 1966). Also, it became increasingly clear that this distinction between competence and performance was not trivial, and that the competence theories proposed by linguists could not simply be transferred to performance. Another difficulty was that linguistic theories were changing rapidly and that made it more difficult for psychologists to test them.

As a result, there was relatively little interaction between the study of psychology and linguistics for the next couple of decades. Psychologists were still interested in language—very much so—but focused more on issues of performance, such as the processes by which syntactic structures are constructed in real time, how ambiguities in language are resolved, and how word knowledge is accessed upon encountering a word. Linguists were still interested in language as a mental phenomenon, but focused on issues of competence—what knowledge of a language entails, and formulating theories that could apply to all languages, regardless of the apparent differences among them.

The separation between these two fields was not to last. Starting in the late 1980s and early 1990s, there was a renewed interest in psycholinguistics as a joint venture between linguists and psychologists, and these days many researchers have been trained in both disciplines (as well as other fields). Substantial advances in related disciplines, including neuroscience, computer science, and cognitive science have given researchers interested in language processing a huge new set of resources, both in terms of knowledge and tools. We now have researchers working on computational models of language processing, informed by current knowledge about how the brain works. We have researchers working at the intersection of language and other cognitive abilities, including not just working memory, but things like scene perception and reasoning. It is actually a very exciting time to be a psycholinguist—both because of this explosion of new interdisciplinary endeavor and because prior work has meant that we have started to build a solid foundation of understanding about how language processing works.

MAJOR THEMES IN PSYCHOLINGUISTICS

There are several key issues, or themes, that we should be aware of when thinking about how people process language. These are actually not limited to language processing, but are related more broadly about how and when information is processed. Nonetheless, they are important to how language is processed and as such they have shaped the major approaches and theories within psycholinguistics. We will discuss each of the most relevant themes in turn in the following.

Top-Down Versus Bottom-Up Processing

This first theme and the next are really about the nature of information flow: do people rely solely on the information from the input that they receive or do they also use information from "higher" levels of processing to understand language. In a strictly bottom-up processing model, only information from the input is considered—processing is entirely stimulus-driven. For example, when you hear a language sound, you must decide what it is. A strictly bottom-up processing model would say that you use information from the sound itself to determine the linguistic identity. For example, to understand the word "file," you would need to process the details of the sound waves as they come in—which frequencies are louder, which are quieter, how long certain frequencies are loud, and the like. A top-down processing model, on the other hand, *adds* additional information from "higher" processes. Let's say that you are listening to a speaker after you have asked her for directions to the library. As she answers, someone honks the horn of their car right next to you and so part of her answer is obscured. The sound stream that reaches your ears is something like "Well, to get to the li[beep]ry, first you need to. . . ." In a bottom-up model, you have a set of sounds that you cannot hear because the horn makes them impossible to hear. In a top-down model, you may use information from your knowledge of sentence structure (the obscured word is a noun) and discourse (we're talking about the library) to help "hear" the missing sounds. So, while you definitely hear the car horn, you can't say exactly which sounds it obscured because you feel that you heard the word *library*.

Clearly, we use bottom-up information when processing language—we need to pay at least some attention to the actual input, otherwise we wouldn't actually be listening to people talking to us. However, in experimental conditions that mimic the real-world example earlier, the results show that people also use "top-down" processing to fill in the blanks—they report hearing sounds that are physically just not present in the speech stream (because the experimenters removed them) or are deliberately ambiguous (because the experimenters made them so), and further, they "hear" sounds that are consistent with interpretation from higher processes (e.g., Warren & Warren, 1970). So, both top-down and bottom-up processes are important. The bigger picture here is whether and to what extent information from later or higher processes can influence early or lower processes. This is a recurring theme because it is applicable not only to how we process sounds, but also to how we build and interpret sentence structures.

Serial Versus Parallel Processing

This second issue (also related to information flow) is whether individual processes related to language need to completely finish before further processing can happen, or whether processes related to the same information can overlap. In serial processing, each step must be finished before the next can begin. In parallel processing, steps may overlap so that one process may not be completely finished before the next one begins. For example, say that we are trying to determine the meaning of a word we're hearing—do we search through a mental list of word entries, considering and rejecting until we get to the word that matches what we're hearing (a serial approach) or do we somehow keep all the possible words in mind at once, considering them all at the same time until one of them becomes the clear winner (a parallel approach). Or, assuming we've gotten the right word, do we need to completely finish the word selection process before turning to placing the word within the sentence structure (serial), or can we begin to build structure before we are completely certain what the word is (parallel)?

Automatic Versus Controlled Processing

Another key issue is whether a particular process is automatic or controlled. In general, psychologists have established that we handle information using a finite set of resources. Some processes tax these resources more than others. Automatic processes are those that do not tax resources very much. In general, they tend to be unintentional, uncontrollable, efficient, and fast. One commonly cited example of an automatic process is our ability to roughly estimate the frequency of events. For example, if you were asked whether red cars are more common than yellow cars, you would be able to easily say that red cars are more frequent, even though you've probably never stood out on a street and made a tally. In terms of language, understanding language sounds is largely done by automatic processes. Controlled processes, in contrast, are those that do require more resources and are slower and can be subject to strategic effects. Imagine trying to learn how to play a new musical instrument—at first, you will need to pay close attention to coordinating your hand movements to produce the sounds. This is a controlled process (at least at first, it can become automatic through practice). Building up the structure of a sentence is also largely a controlled process.

Modularity

The final key theme is the issue of modularity. This actually comes in two flavors: first, there is the degree to which individual processes within language processing are distinct and isolated from each other and second, there is the degree to which language as a system is distinct and isolated from other cognitive systems. With respect to modularity generally, Fodor (1983) proposed several key properties of modular (isolated) processes: modular processes are specific to a particular domain (they are not spread across multiple areas), automatic, fast, and not affected by feedback from other processes. From the perspective of individual processes within language, this then raises the question of how much different processes interact. For example, is it possible for processes related to connecting words in sentences to interact with the semantic processes necessary to understand them? From the perspective of broader cognitive processes, there is also an outstanding question concerning the extent to which language is independent from (or modular with respect to) other cognitive processes.

A QUICK OVERVIEW OF THE REST OF THE BOOK

In this book, we will explore a set of key topics that have shaped research and given us a much better understanding of how language processing works. First, however, we will go over a brief introduction to language in chapter 2. As we have already seen, the study of language involves examining sounds, structure, and meaning, and in this chapter we will cover the aspects of language in each of these areas that are most relevant to psycholinguistics. We will discuss the sounds of language and how different languages may group and order these sounds, and we will also examine the internal structures of both words and sentences. In this chapter, it will become clear that while language is complex, it also follows rules that we can define and use to predict what is acceptable or unacceptable (grammatical or ungrammatical) in the language. For example, in English we could invent a new word like “pilt” but not like “tlip”—in chapter 2 we will examine why this is the case.

In chapter 3, we will discuss some of the clever techniques that researchers in psycholinguistics use to uncover what are largely unconscious processes. As we shall see, researchers need techniques that provide a window into processes that occur on the order of milliseconds, and that can probe language processing in a way that is still valid even when people know that you are studying how they produce or comprehend language. We will cover relatively low-tech methods that simply involve pencil and paper as well as very high-tech methods like functional magnetic resonance imaging (fMRI) that use advanced technology to determine brain activity in response to language. We will also discuss some of the techniques that researchers use to ensure that their experiments test what they think they are testing and that the participants in these experiments can’t guess what the experiment is about.

In chapter 4, we will discuss a topic that has dominated the field for over two decades—how people handle ambiguity in language. This is a huge topic in part because so much of language is ambiguous and so understanding language processing necessarily means understanding how people are able to process ambiguous language. However, as we shall see, ambiguity resolution is also at the heart of one of the key themes in psycholinguistics—information flow. As we encounter sounds and build them into words and then into sentences, what information are we able to use to guide this process? Are we only able to use current, incoming information—the sounds themselves, the words as they are selected, and so on. Or, are we able to use information about the larger context of the sounds and words, and if so, which aspects of that context? Can real-world knowledge, or prior semantic content in a sentence, change how we interpret the structure of the sentence, or the meanings of the words within it? For example, if you hear a sentence like “Bob took his ATM card and went to the bank,” do you—even for a moment—interpret “bank” as a piece of land next to a body of water?

In chapter 5, we will discuss how language is represented, both in the brain itself and in how multiple languages interact. We will cover which parts of the brain are critical for the basics of language, and how language ability can be disrupted when the brain is damaged. We will see that there are different ways in which language function can be impaired, for example, having difficulty with understanding and producing the structure of language without difficulty in understanding the words themselves. We will also talk about progressive language disorders like semantic dementia and what the study of disordered language can tell us about the neurological basis of language. We will also discuss how language is represented when someone speaks more than one. Some key issues for bilingual speakers are how knowing multiple languages impacts language representation in the brain, and to what extent the languages interact both with each other and with other aspects of cognition.

In chapter 6, we will take a look at research on how language works in a broader context—including how speakers and hearers influence each other in a dialogue and how prior context influences how we process language. While speakers clearly influence listeners in terms of telling them new things, there can also be more subtle influences from the speaker that can make language processing easier or harder for the listener. Research has also looked at what impact listeners have on their speakers, and how participants in a dialogue coordinate their language over time. In this chapter, we will also address about how context influences how people refer to things—for example, if you wanted to talk about your pet, what determines whether you would use a description (like *the cat*), a name (like *Fluffy*), or a pronoun (like *it*)? Further, what influence would your choice of wording have on the person you were talking to?

In chapter 7, we will look at sign language research to see if and how sign language processing differs from speech. This has very important implications for how we think language is organized because it can tell us which parts of language processing are due to language as an abstract, symbolic system, and which are specific to the modality—the form of the language, spoken or signed.

Finally, in chapter 8 we will look at a relatively new hypothesis that has emerged: most previous work has taken for granted that comprehenders (and speakers) fully process language, that is—that we try to build complete representations of what we hear, read, or produce. However, what if this is not the case? What if, instead, we often only build an incomplete representation that is merely "good enough" for the linguistic task at hand? The answers to these questions have the potential to radically change how we think about language processing.

REFERENCES

Blumenthal, A. L. (1987). The emergence of psycholinguistics. *Synthese, 72*, 313–323.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Miller, G. A., & McKean, K. E. (1964). A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology, 16*, 297–308.

Seuren, P. (1998). *Western linguistics: An historical introduction*. Oxford, UK: Blackwell.

Skinner, B. F. (1957). *Verbal behavior.* Acton, MA: Copley Publishing Group.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior, 5*, 219–227.

Warren, R. M., & Warren, R. P. (1970). Auditory illusions and confusions. *Scientific American, 223*, 30–36.

Wundt, W. (1900). Die Sprache, In *Völkerpsychologie*, Vol. 1 & 2. Leipzig: Alfred Kröner Verlag.

2

Language as an Object of (Psychological) Study

In this book, we're going to talk a lot about language, and so it makes sense to start with an overview of some language basics, particularly those that are relevant for the rest of the book. Language is a complex but rule-governed system with roughly 6,900 currently spoken versions that are, by definition, mutually unintelligible. Some languages, like English, Hindi, and Japanese, have millions of speakers. Other languages, like Yukaghir (spoken in Siberia) and Pawnee (spoken in the United States), have fewer than 100 speakers (473 languages are considered "nearly extinct"; Lewis, 2009). Some languages have clearly defined words that get combined into sentences, whereas other languages have long, single words that carry the same meaning as an entire sentence. The sheer number, differences, and complexity of language, combined with the relative recentness of modern linguistic study, have meant that we still do not have a complete account of how every individual language works, nor a complete, unifying theory of how language as a whole works. But, we have some pretty good ideas.

While anything like a complete account of the features of language is beyond the scope of this book, interested readers can refer to any introductory linguistics textbook for more information on any of the following topics (*Language Files* provides a good introduction), and there are hundreds of specialist books on each of these topics, including in different languages.

BASIC FEATURES OF LANGUAGE

The study of the properties of language can be divided up into roughly five, somewhat overlapping categories: sound system, word structure, sentence structure, meaning, and real-world use. We will focus primarily on the first three because it is in these cases that the insights from linguistics have been directly relevant to psycholinguistics. However, we will also briefly discuss meaning and real-world use and the major issues to consider in these areas.

Sound System

One of the hallmarks of language is the way in which units of meaning (like words) are composed of smaller, nonmeaningful segments. These segments can be combined and recombined in a potentially infinite number of ways, allowing for a potentially infinite number of words. In spoken languages, segments are sounds—each language has a set of sounds that are produced by changing the positions of various parts of the vocal tract. There are two important aspects to these segments: first, they have no independent meaning. So, there is nothing about the "b" sound in the word "bird" that itself means "bird"—the sounds are arbitrary. Other languages use different sounds to refer to the same thing. The reason that this is important is that it makes language very flexible, so we can create new words for things using any sounds in our language. The second important thing to know about sound segments is that they cannot be combined in any old way—they are governed by rules that are specific to each language. These are not eighth grade English grammar rules that must be memorized and have a thousand exceptions. These are rules that one knows unconsciously by virtue of knowing the language and, almost by definition, don't have random exceptions. For example, in English it is possible to have the sounds "t" and "l" together, like in the word "little." However, it is not possible in English for this sound combination to come at the beginning of a word—"tlip" is not a valid word in English, though it could be in another language. This is important because it means that even at the level of sound combinations, language is a rule-governed system. However, complex it might seem at times—it is not random.

The sound system of language is actually studied in two main parts: *phonetics* is the study of the actual sounds in a language and how they are produced, whereas *phonology* is concerned with how sounds are mentally categorized and the rules that govern how they are combined.

Phonetics

We can define the sounds of language in terms of how they are produced, or *articulated*. We can divide sounds roughly between vowels and consonants, and each of these categories can be further divided according to specific parameters. For consonants, we can define most individual sounds according to three main parameters: place of articulation, manner of articulation, and voicing. Place of articulation refers to which parts of the vocal tract (i.e., the throat, mouth, and nasal passages) are involved in making the sound—for example, a "b" sound is made using our two lips (a "bilabial" place), whereas "d" is made by placing the tongue against the ridge right behind the top front teeth (the *alveolar* ridge, and the so-called "alveolar" place of articulation). Manner of articulation refers, more or less, to the degree of airflow restriction—whether the air is stopped completely, as in "b" and "d" (called a stop or "plosive" manner of articulation) or whether it is only highly restricted, such as in the sound "z" (called a fricative manner). The final parameter, voicing, refers to whether the vocal folds in the larynx (the "voice box") are vibrating or not. When these folds are pulled shut and air is forced between them, this causes a vibration—voicing. The folds can also be pulled open, in which case they don't vibrate and this produces a voiceless sound (there are stages in between, too, such as whispering). The sounds "s" and "z" in words like "sue" and "zoo" differ only in whether the vocal folds are vibrating or not—that is, they differ in their voicing.

Vowels are somewhat different from consonants. First of all, while consonants may be voiced or voiceless, vowels are voiced. Also, the airflow through the vocal tract is not constricted the way that it can be for consonants (there are no "stop vowels" or "fricative vowels"). Different vowels are produced by changing the position of the tongue in the mouth and rounding (or not rounding) the lips. Vowels tend to be louder and longer than consonants.

Not every language makes use of every possible sound. For example, English does not have a "bilabial fricative," although several other languages do. Conversely, English has a dental fricative (the "th" sound in words like "teeth" [voiceless] and "teethe" [voiced]) that many other languages do not. This poses a problem for writing down language sounds because it means that there is no language with a writing system that has symbols for every possible human language sound. To make matters more complicated, some languages (like English) have a single letter for multiple sounds (e.g., the "c" in *species* and *cat*) or use multiple letters or combinations of letters for a single sound (e.g., the "k" sound in *cat*, *pack*, and *park*). So, linguists and other language researchers have developed a special alphabet, the International Phonetic Alphabet (IPA), to transcribe sounds in languages. Many of the symbols are taken from the regular

Latin alphabet (what you're looking at right now), but there are lots of other symbols, too. For example, the sounds that we spell as "th" (e.g., tee*th* vs. tee*th*e) are transcribed as [θ] and [ð] so that the word that we write as *teeth* in IPA would be [tiθ]. According to the IPA, there are approximately 12 distinct places of articulation (depending on how you count them) and 9 main manners of articulation. However, there are some other parameters that can be applied to consonants, and in the end there are over a 100 possible consonants available for language use. There are fewer vowels, with about 25 different possibilities.

In addition to transcribing sounds using a special alphabet, the other way that one could write down or visualize the sounds that we produce for language is to show what the sound waves of the sound look like. These are called spectrograms, and provide a way of looking at how the small changes that we make in our vocal tract impact the actual sounds that we produce. Spectograms basically show how loud (or quiet) all the different frequencies in the human voice are as speech unfolds. For example, for vowels, certain bands of frequencies become notably louder than others, and we can use the relative spacing between these bands (called formants) to help identify the vowel.

Phonology

Describing the way language sounds work in the section above makes the whole system sound fairly straightforward, but there are in fact a number of complicating issues with understanding speech sounds. Crucially, we need to get from sound waves to language sounds. Phonetics is principally concerned with the physical sounds of language, whereas phonology is concerned with the rules that different languages apply to these sounds, including how the physical sounds map onto linguistic categories of sound for a given language.

Languages often group together similar sounds into a single sound category—called a phoneme. To better understand this—let's start with some example words: "little," "kitten," "top," and "stop." Most speakers of English, when asked if there is a common sound that is shared by all these words will say yes—it's the "t" sound. However, in fact, the sound that English speakers hear as "t" in these examples is actually different in each of the words, regardless of speaker. In "little," the "t" sound is formed by briefly tapping the tip of the tongue against the alveolar ridge. In "kitten," the "t" is formed by briefly constricting airflow at the voice box (called the larynx). The "t" in spot is more similar to the "t" in little, but importantly differs from the "t" in the very similar "stop" because it occurs with a short puff of air at the end. This seems like such a trivial difference, but in some languages (like Hindi), the difference between a "t" without a puff of air and a "t" with a little puff of air at the end is as different to listeners as the difference between "t" and "d" is to English speakers. These different sounds that all form a single phoneme are called allophones. Thus, in some sense, a phoneme is like a category of similar sounds. So, part of knowing a language is being able to immediately and correctly place allophones into their phonemic category. In other words, we need to be able to identify a sound and know what category of other similar sounds it gets grouped with.

As mentioned earlier, another important aspect of sound in language is that languages will typically have rules about where phonemes can occur and what other phonemes they can or can't occur with. Additionally, languages have rules about how phonemes will change (in terms of how they sound) depending on the sounds next to them. There are rules that cause sounds to be added, deleted, or changed, depending on the "environment" that they occur in. For example, English has a rule that applies when we try to add an "s" sound to make a word like "fox" plural: we insert a vowel between the final consonant "x" and the plural "s"—and in this case we actually reflect that new vowel in the spelling—"foxes." Other changes, even additions and deletions, are often not reflected in the spelling of a language and go relatively unnoticed by speakers.

Another interesting sound feature to make matters more complicated is that some languages use tone—the pitch with which a word is said—to make meaningful distinctions between words. So, the same set of sounds said at one pitch will mean one thing, but the same set of sounds at another pitch will mean something entirely different. This means that listeners must not only decipher the individual sounds, or segments, in a word, but must also, depending on the language, pay attention to the pitch with which they are said. English is not a "tone language," but we do use stress in a similar way. For example, if you read the following sentence aloud, "Bob wanted to record his voice on a record," you'll notice that while you say "record" twice, each version is somewhat different. When it is a verb, "record" is said with stress on the last part of the word "reCORD," while when it is a noun, the stress comes on the first part, "REcord."

With respect to psycholinguistics, big issues related to the sound system of languages are how sounds are actually produced and how we are able to correctly identify speech sounds in real time. This is actually not a trivial process. In addition to needing to apply the sound rules of the language (i.e., recognizing allophones as phonemes and compensating for rule-based additions, deletions, and changes), listeners have to face the fact that speakers rarely produce the same sound in exactly the same way—each time we produce a sound it is slightly different based on things like the other sounds next to it, the stress placed on the word or syllable with it, even the emotion of the speaker. And, because speakers vary in size and shape, no two speakers produce exactly the same sound—the spectrograms never look quite alike. So, listeners must not just hear the sound segments, but be able to interpret them correctly even though there is not a single cue in the sound that will always tell them what the sound is.

Words and Word Structure

Phonemes can be combined (and recombined) to make words, and words themselves have an internal structure and can even be ambiguous based on this structure. For example, does the word "unlockable" mean "able to be unlocked" or "not able to be locked"? It can actually mean both, and it depends on how you build the internal structure of the word. Morphology is the study of the structure of words—what rules guide how they can and can't be built. Many words are made of smaller, meaningful parts that aren't actually words—things like "un" and "ness"—these are called morphemes. Morphemes are any sets of sounds that carry meaning. So, while suffixes and prefixes are morphemes, so are whole words, like *cat* and *dog*.

Languages vary a lot with respect to how they allow words to be built and linguists have identified a number of different types of languages with respect to their morphology. In some, like Mandarin Chinese, most words are composed of a single morpheme—these are sometimes called "isolating" languages because each morpheme is isolated. Others languages, like Sora (spoken in India), have words that incorporate lots and lots of morphemes. In some of these types of languages, words can be very long, and actually encode meaning that corresponds with a whole phrase in another language. Some languages have a lot of morphemes that get added by more or less sticking a morpheme onto the beginning or end of another morpheme. In these "agglutinating" languages, it is relatively clear what all the individual morphemes are that make up a word. In other languages, morphemes may blend together more and obscure the distinction between one morpheme and another. Further, linguists have identified different types of word-building rules that languages may use. For example, one type of rule that is familiar to speakers of English is adding a nonword morpheme to the end of another morpheme—called "suffixation." A rule that is not so familiar to English speakers is "reduplication," in which one morpheme is repeated. So, for example, in English, we make a word plural by suffixation—adding an "s" to the end of the thing we want to make plural—but in Indonesian, the plural is formed by repeating part of the word. For example, the singular form of the word house is "rumah" but the plural is "rumahrumah."

While words are clearly formed using rules of various types, it is not the case that every word may follow the *same* rule. Notice that even setting aside differences in phonology that we discussed earlier, English doesn't always form the plural in the same way. We've actually got several plural forms. Some words like "child" and "ox" are made plural with another suffix (-(r)en). Others are made plural by changing the vowel (man to men), while still others don't actually appear to change, for example, deer and deer. Yet, by and large it appears that adding "s" is the "default" rule in English: new words in the language are most likely to have "s" added to make a plural version. There are important and interesting historical reasons for how this situation has come to be, but the key implication that has driven a lot of research on word formation is that we effectively have two systems—a rule-based system in which plural is formed by adding "s" and then a set of historical exceptions, which are stored whole in our mental dictionary (called the lexicon). This is a controversial idea, but leads us to one of the big questions for language processing: Do people build words from morphemes and word-building rules in real time as they speak and/or hear them, or are words stored as whole units in our mental dictionaries? So, for example, when you hear "cats," do you look up "cat" and "s" in your lexicon, or do you have an entry for *cats* that is separate from the entry for *cat*? The short answer appears to be that we do both—it depends on the frequency of the word and the types of morphemes involved. A number of researchers have also successfully shown computationally that one doesn't need to divide the lexicon strictly between rules and exceptions to those rules in order to account for how language appears to work. Instead, it may well be that our word processing system is sensitive to regularities of all kinds between the forms of different words, which would allow us to account for the fact that while people *usually* apply the apparent default rule to a new word, if the word resembles an irregular word, people may use the "exceptional" rule: the regular past tense in English is formed by adding "-ed" to the end of a word, but if you give people a new word like "fring" and ask them to make it past tense, they may say "fringed"—as in "Yesterday, I had fringed." But, they may also say "frung" instead because “fring” sounds a lot like some other words “sing” and “bring,” that don’t add “-ed” for the past, but rather change the vowel (e.g., “Yesterday, I had sung/brung…”).

Sentence Structure

Words, including their internal sounds and structure, are crucial to languages, but language doesn’t occur as just a set of isolated words, but instead as words put together in systematic, rule-governed ways. Further, at this level language is built in a hierarchical way such that words can be grouped together into phrases. Syntax is the study of how sentences are formed. This is an important part of language for psycholinguistic researchers, because it poses many questions about how people understand these groupings of words.

Rules for Creating Phrases

Rules for creating phrases generally apply at the level of types of words rather than particular individual words. So, rules apply to things like nouns and verbs, not to particular words like *dog* and *eat*. However, properties of individual words can and do play an important role in sentence processing as well.

We can break any sentence down in smaller, connected phrases. For example, we can take a sentence like this and break it down:

1. The artist bought a paintbrush.

First, there are two noun phrases (NPs) in the sentence—*the artist* and *a paintbrush*. There are rules that establish how a phrase can and can’t be made in a particular language. In English, one part of the rule for making noun phrases is that the article (also called a determiner) comes before the noun. Importantly for English, there is no rule that allows the article to follow its noun—it is not grammatical in English to say *Artist the bought paintbrush a*, even though in other languages it can be.

Next, we can create a verb phrase (VP) by combining the verb *bought* with the noun phrase that follows it: *bought a paintbrush*. Here again, the order matters in English and it would not be grammatical to put this noun phrase *before* the verb. In English, we can't say *The artist a paintbrush brought*. However, again, this could be (and is) grammatical in other languages.

This brings up an important point—language researchers are generally not interested in how people "should" speak, but rather how they actually do. And so, the rules of syntax or grammar for language researchers are not about style (e.g., whether you should end a sentence with a preposition or not), but rather are intended to capture the facts about how people actually speak. Once again, what linguists seek to understand is not your eighth grade English teacher's grammar.

Grouping and Linking Phrases

As with sounds of language, something that initially seems pretty straightforward gets complicated fairly quickly. First, it is not always clear how to combine or group phrases—sometimes it's ambiguous, like in the following example:

2. The policeman watched the tourist with the binoculars.

Who has the binoculars? It's not clear from the sentence—they could belong to the tourist or the policeman could be using them to do the watching. This ambiguity means that there are actually two possible structures for the sentence, one in which the binoculars are grouped with the tourist, and one in which they are grouped with the verb *watched*. A big question for language processing is how people decide which interpretation to pick as they hear or read the sentence. We will explore this question in detail in chapter 4.

Another puzzle is how people are able to link up phrases even when they are not right next to each other. In fact, this brings up an obvious question—why *not* just think of sentences as words strung together—why use phrases at all? There are actually lots and lots of reasons to think that sentences are organized by phrases—books and books worth, in fact. Let's just look quickly at one example that shows why we need phrases:

There is a rule in English that the subject of a sentence and its verb are linked such that the form of the verb changes to "agree" in number with its subject. For example:

3. The artist **wants** to borrow some money from Mary.
4. The artists **want** to borrow some money from Mary.

The thing to notice here is that the verb *want* either has an *s* or not, and that what determines when the verb should have an *s* is whether the subject *artist(s)* is singular or plural. For example, the following sentences are ungrammatical in English because the verb "disagrees" in number. The standard convention in linguistics and psycholinguistics is to mark ungrammatical sentences with an asterisk at the beginning:

5. *The artist want to borrow some money from Mary.
6. *The artists wants to borrow some money from Mary.

But, this isn't enough to convince anyone that we need phrases. So, what happens when the subject gets a little more complicated?

7. The artist that Mary once met at a party wants to borrow some money from her.

Here, the verb *want* still needs to agree in number with *artist*, not with *Mary* or *party*. Of course, all of these nouns are singular, and so it's actually not clear just from this example whether this is true, but we can look at other versions of this sentence:

8. The artists that Mary once met at a party want to borrow some money from her.

9. *The artist that Mary once met at a party want to borrow some money from her.
10. *The artists that Mary once met at a party wants to borrow some money from her.

By looking at all four versions of this sentence, we can see that the same rule applies, and it's the main noun (called the head noun) of the phrase that counts, not any of the intervening nouns, not even the one that occurs right before the verb. So, the verb is not "agreeing" with the noun closest to it, but with the noun that is in charge of the whole noun phrase.

The key questions for language processing are how the words and phrases in a sentence are linked together in a meaningful way, and how the structure of the sentence gets built, both by speakers and by comprehenders.

Meaning

The field of semantics is concerned with meaning in language and can be divided into two major parts: lexical and propositional. Lexical semantics focuses on word meanings—for example, the meaning of the word "cat." Propositional semantics focuses on the meanings of sentences—how we know who did what when we encounter a sentence like "the cat sat on the mat." In both cases, there has been a debate about what language refers to. Does it refer to the real world or to representations of the real world that we have in our minds? Psycholinguistics has adopted the latter view—words (and sentences) refer to representations of the real world that we carry in our heads. This means that in order to understand meaning in language from a psycholinguistic perspective, we need to know what these mental representations are like and how language links to them.

In this book, we will focus largely on lexical semantics, addressing the nature of meanings of concepts and their links to language in chapter 5. But, we will also consider how the meanings of sentences and individual words may be linked to larger, dynamic representations of situations in chapter 6. Another big question in lexical semantics that we will discuss is how people get to the right meaning of a word when there is more than one—if someone talks about going to a bank, how do we know whether they are headed for a financial institution or down to the river. We will address this in chapter 4.

Real-World Use

Of course, language is generally used in a larger context with some sort of purpose in mind. There is a great deal of linguistic study of the social and functional aspects of language use that have not been explored from a psycholinguistic perspective, but one area of study that has gotten an increasing amount of attention in recent years is the form of the words that we use to refer to things. For example, when someone is talking about, say, a penguin, they might first introduce the penguin in exactly that way—that is, *a penguin*, but subsequent references to the penguin would be different—for example, we could use "the" (i.e., *the penguin*), or "the" plus a category (i.e., *the bird*) or a pronoun like "it" (or even "he" or "she"). As it turns out, speakers are actually quite systematic in terms of the form of reference that they select and a sentence can sound downright odd when the "wrong" form is used. Imagine for a moment that a friend walks up to you and, after saying hello, starts with "After I fed those fish to that penguin. . . ." Unless you know that this particular friend works with penguins, your quite natural response might be "Huh? What fish? Which penguin?" Your friend would sound downright odd.

Notice here that while for phonology, morphology, and syntax we could talk about rules that might be broken and thus result in ill-formed utterances, in topics related to real-world use, there are fewer iron-clad rules and so sentences and words may be perfectly well-formed and even sensible, and yet not be appropriate or felicitous for the context. And so, in this area there is an additional complication: we must consider the context of the language—not just prior linguistic context (i.e., what people have just said or written) but any information that is shared among the speakers/listeners—including mutual knowledge between a listener and speaker that is derived from knowledge shared by virtue of being members of the same community (Clark & Marshall, 1981). For example, if your friend walked up to you while you were both visiting an aquarium and you were standing next to a set of penguins and a bucket of fish, then the statement above might actually be felicitous.

Linguists have been interested in accounting for how language refers to the same thing in different contexts—in other words, what determines the form of referring expressions. Gundel, Hedberg, and Zacharski (1993) proposed the Givenness Hierarchy, which is a set of cognitive statuses that can be matched to the form of referring expressions. In this approach, the use of different determiners and pronominal forms signals different mental statuses and restricts the set of possible referents of the referring expression for the listener. This hierarchy is shown in the following table:

Status	**Form**
In focus	*it, he, she, they*
Activated	*that, this, this* [noun] (definite) (e.g., *Look at this rabbit right here.*)
Familiar	*that* [noun] (e.g., *You could get that rabbit you saw the other day.*)
Uniquely identifiable	*the* [noun] (e.g., *The child picked up the rabbit.*)
Referential	*this* [noun] (indefinite) (e.g., *While I was walking down the street this rabbit just hopped across the path.*)
Type identifiable	a [noun] (e.g., *A rabbit ate my lettuce!*)

The relationship between the statuses is such that if a referent is, for example, uniquely identifiable, it is also referential and type identifiable. But what do these mean? For a referent to be *type identifiable*, the listener must be able to identify a representation of the type of object that is indicated by the reference. So, in a sentence like, "A rabbit ate my lettuce!" the speaker can assume that the listener knows the meaning of "rabbit" and can know the *type* of thing it refers to. The scale progresses up in order of how identifiable (given) a referent is until something is *in focus*. This status is actually referred to quite a bit and tends to mean slightly different things to different researchers, but in essence what this means that is pronouns are most often used to refer to things that are at the current center of attention. As we shall see in chapter 6, how pronouns are interpreted during processing is actually somewhat complicated, but this basic idea is still essentially true.

Another, similar way that this idea has been framed is in terms of the accessibility of the referent (Ariel, 1988, 1990). In this case, it is the degree of activation or "accessibility" of a referent in a mental representation that determines the form of its referent. Pronouns will mark the high accessibility of their referent, which means that the referent will likely: (1) have been mentioned within the same sentence, (2) have few alternative referent possibilities, (3) be in cognitive focus, and (4) be mentioned in previous discourse. In this approach, low accessibility markers (like definite descriptors) usually refer to referents that are stored in long-term memory (which corresponds to the mutual knowledge shared as member of the same community). High accessibility markers like pronouns, on the other hand, refer to referents in short-term memory.

Both of these approaches, and others, have thus centered on the idea that for reference and coreference, the status of the referent is crucial to the form of the referring expression. This status is, in turn, dependent on all kinds of linguistic and real-world factors, and so it is in this type of study that we see how language and nonlinguistic factors are linked.

REFERENCES

Ariel, M. (1990). *Accessing noun–phrase antecedents*. London: Routledge.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics, 24*, 65–87.

Clark, H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). New York: Cambridge University Press.

Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*(2), 274–307.

Lewis, M. P. (Ed.). (2009). *Ethnologue: Languages of the world*. Dallas: SIL International.

3

How We Know What We Know: Methods in Psycholinguistics

While physics and chemistry are referred to as the "hard" sciences and psychology and similar fields are considered "soft" sciences, it is in fact quite hard to study almost any aspect of human cognition. And, unlike some other aspects of our wondrous human abilities, language is generally considered to be unique to us. That is, because the difference between, say, human visual abilities and those of other mammals is really one of degree, we can use animals instead of humans as a subject of research when the going gets sticky (and skulls need to be opened). However, the difference between human language and other animal communication systems is considered by many to be one of *kind*—a qualitative difference rather than merely a quantitative one. It is not the case that humans just have some kind of enhanced version of an animal communication system. (In fact, some researchers argue that we still have the remains of a shared animal communication alongside language.) With language, we really do seem to be doing something different. This means that it doesn't do us as much good to turn to animals for help with language research; laboratory rats just aren't going to be able to tell us anything about how ambiguous words are processed, or how we understand answers to questions. Further, because we aren't willing to crack open people's heads just to have a look, we need to find ways to study language in people without resorting to invasive surgery.

So, we can't rely on the same "objective" observations that our colleagues over in the physics department use. For example, even if we could take a bit of brain and measure, prod, and (gently!) electrocute it without invasive procedures, this would still not yield data that allow us to say that we've figured out something like relative clause processing. To make matters worse, language is one of those things that we do without consciously thinking about it—that we *have* to do without needing to think about it. Certainly, we can choose our words carefully in a debate, or when caught eating the brownie that we promised we'd save. But, even then we're mostly thinking about the content of what we want to say—we're not thinking "maybe I should stick in a relative clause here.... oh! don't forget the determiner!...did I need an inflection on that last verb?...," and so on. Try asking someone how they know what "The horse raced past the barn" means. Go ahead. Or even, "The boy kissed the girl." It doesn't work—we process language on the order of milliseconds and we simply don't have conscious access to the action behind the scenes. In short, we can't just ask people how they do it. We in the "soft" sciences are stuck with the hard job of doing science on something that's hard to see—a moving target, through a glass darkly.

And this leads us to the next problem: simply put, rocks don't try to second guess you. When looking at natural phenomena in the universe, the universe does not stare back.

And, while the basis of Heisenberg's uncertainty principle is that observation changes what you look at, so far as we know, unlike humans, molecular carbon chains aren't trying to divine the purpose of our observation and perform differently as a result.

So where does that leave us? Needing to be really, really clever. And that is exactly what language researchers have been. In first half of this chapter, we're going to cover an overview of the techniques that are used to measure language processing. In the second half, we'll have a look at the things psycholinguists do when designing experiments in order to ensure that their results are valid.

KINDS OF RESPONSES—WHAT CAN WE MEASURE?

In this section, we'll cover a wide variety of different things that researchers can measure to learn more about language processing. We discuss these techniques roughly in order of how fine-grained a view they provide of the time course of processing (sometimes called temporal resolution)—from techniques that provide information about the end result of processing (low temporal resolution) to techniques that give us an almost millisecond-by-millisecond view of language processing in real time (high temporal resolution). There has been a definite bias, especially in recent years, toward the high temporal resolution side of things. So much so that Mitchell (2004) felt the need to defend the continued use of "lesser" techniques. In fact, while some research techniques do give us a pretty extraordinary window into human information processing as it happens (not to mention pretty pictures of brains, in some cases), there is no technique, or even set of techniques, that is the champion. Each technique has strengths and weaknesses, questions that it is particularly suited to answer and questions it cannot answer.

Think of it this way: if you want to paint a portrait of your favorite pet, you're going to need a set of small brushes. However, if you want to paint a room in your house, then it would be time-consuming (and pretty ridiculous looking) to use those same brushes—you'd need to use bigger ones. And those bigger brushes would not be so good at painting portraits. One can think of this as the notion of "the right tool for the right job" applied to science. A good researcher is going to have a collection of brushes that can handle whatever project the researcher has. And very good researchers carefully match the brush to the project.

OFF-LINE VERSUS ONLINE MEASURES

Another way to think of temporal resolution is in terms of measuring off-line versus online responses. Off-line is the term given to measures that provide information about the end state of processing—when the time course of processing is not something available. The limitation of the techniques that produce off-line measures is that researchers can't isolate which processes are responsible for the data. For example, if you ask someone to respond in some way at the end of a sentence, all kinds of things have already happened—processes related to retrieving word meaning, building and interpreting structure, understanding who did what to whom, integrating information with previous knowledge, just to name a few big ones. So, with off-line measures it is simply not possible to tell which part of the language process contributed to the response, although with careful designs you can do some general isolation. But these are not techniques that get used when the researchers want to know exactly when something happened, or exactly which process is the source of some effect. Instead, these measures get used when researchers are interested in broader questions about processing, such as how people interpreted the meaning of sentence, or how it has influenced their internal (mental) model of the state of affairs given by what they've just heard or read.

Online measures include any measure considered to give information about language processing as it happens. For many researchers, "online" is a relative term—some measures are very "online," whereas others are less so. One way to think of it is to imagine off-line and online as end points on a continuum rather than as strict categories. What makes a measure less "online"? It depends somewhat on the technique, but, in general, techniques that require interrupting the participant's language processing are considered less online, because they risk disrupting the process altogether.

TASKS THAT COLLECT DATA FROM CONSCIOUS RESPONSES

Questionnaires

The prototypical off-line measure is the questionnaire—literally asking people for their judgments about what they've just encountered. This is actually a useful and inexpensive technique that can be used to help design stimuli for another experiment, or can be used to compare the final result of processing with the time course of the processes that get you there. For example, say you are interested in how people know what words refer to while reading. What kind of thing comes to mind when you encounter the word cat? If there is a story about a cat, each time "cat" is mentioned, how do you know if the speaker is talking about the same cat or a new one? A researcher may ultimately be interested in things that happen right when someone encounters a word, but they may also want to know what the final interpretation of the sentence or passage might be. So, they could give a questionnaire to people that lists a bunch of sentences or even short passages with certain words underlined and ask people to read each sentence/passage and circle the word that the underlined word refers to. An example is given in the following:

> Sarah admired Susan because she knew certain things.

So, do people, on average, circle Sarah or Susan to indicate that they interpret one of these women as the *she* in the second part of the sentence? Questionnaire studies can tell us this, and researchers can change small things about the sentence to see what influences the interpretation. For example, what if Sarah admired Susan *even though* she knew certain things?

In fact, all kinds of data can be collected from questionnaire studies. Including, for example, how related two words are to each other, or how acceptable people find a sentence. This can be done by using Likert scales, which involve asking people to circle a number on a scale. You may have encountered this type of question in surveys—for example, restaurant feedback cards often ask patrons to rate how good the food was, how clean the establishment was, and so on, on a numerical scale in which one end (e.g., 1) corresponds to "Poor" and the other end (e.g., 7) corresponds to "Excellent." Likert scales can be (and are) applied to all kinds of research questions. To pick just one of many, many examples, in language research, one could ask research participants to circle how typical a particular thing is for a particular category—so on a scale of 1 (very typical) to 7 (very atypical), how typical is a penguin as a kind of bird? Chances are in this case most people would circle 6 or 7—penguins are not very typical birds.

Button Presses

Our first somewhat technological technique is the button press. Unless the time it takes to press a button in response to a stimulus is recorded, this technique is generally just a fancy version of a questionnaire. However, it's possible to record how quickly someone responds with a button press, and in this case a new kind of information is gathered—how long it takes the human processing system to make a decision to respond. This is still toward the "off-line" end of the scale, as are all of the tasks that collect data from conscious responses. However, some applications of it are considered more "online" than others.

The button press task is perhaps the most versatile of all the things that we can do to collect data involving response times. The possibilities are nearly endless. First, we can ask participants to press a button in response to any kind of question or judgment. We can ask them to judge whether a given letter string (e.g., THRIP) is a real word or not. We can ask people to press a button as they read each word of a sentence (more on this in a moment). We can even ask them to keep pressing one button as they read, but switch to a different button on the word where a sentence stops making any sense to them. We can ask them to answer yes/no to questions that quiz their comprehension of what they read.

One application of the button press warrants a little more discussion, and the task that involves it has its own name (and acronym): self-paced reading (SPR). As the name suggests, participants are asked to read at their own pace. The most basic type of SPR works like this: participants are asked to read silently anything that they see on the screen. Let's say the experimenter is interested in how people process sentences (a common application area for SPR), and so the experiment is going to involve showing sentences to the participant. If the experimenter is only interested in how long it takes to read the entire sentence, he/she may display the entire sentence all at once. The participant reads the sentence silently, then presses a button as soon as they have finished. Voila—we have a data point: how long it took between when the sentence appeared on the screen and when the participant pressed the button. Usually, however, participants don't read the sentence all at once, but see it only in small chunks so that the sentence is slowly revealed word-by-word or phrase-by-phrase, controlled (paced, if you will) by the participant's button presses. One particular implementation of this type of method uses dashes to indicate the length of the upcoming words and give the general shape of the sentences that will be read. For example, a participant in an SPR study might, for a particular trial, see the following:

----- ------- ----- ------- --- ---- ------- ------.

Then, when they press a response button, the first word is revealed:

Sarah ------- ----- ------- --- ---- ------- ------.

Then, when they press the button again, the next word is revealed (and the first word is turned back into dashes).

----- admired ----- ------- --- ---- ------- ------.

The participant proceeds to read the entire sentence in this way, and then would move on to read the sentences of the following trials in the same way. There are variations on this—some studies do not use dashes at all; some studies lump together phrases instead of presented things just one word at a time; some studies leave previous words up on the screen; and so on.

SPR is sometimes, mostly jokingly, referred to as "poor man's eye-tracking" and it certainly has some limitations that eye-tracking (discussed later) can avoid. However, it has been an incredibly useful measure over the years, providing a huge amount of important and informative data about how we process written language (especially at the sentence level), and it will undoubtedly continue to do so.

Vocal Responses

The last of the conscious responses we'll talk about here is vocal response. As the name suggests, this applies to cases where the participants verbalize their response. This type of response method is used not only in experiments that are interested in language production (the obvious thing) but also in experiments concerned with things related to comprehension as well. Like button presses, this is a response that can be applied to all kinds of questions. And, just like button presses there are two kinds of data that can be gotten: the nature of the response (i.e., what the participant actually says) and how long the response takes. Unlike button presses, though, in which the only timing data is really just how quickly someone presses a button, vocal response data can provide other timing data. In addition to onset (i.e., when the vocal response starts, the thing that is similar to button press data), we can analyze the duration of the utterance, or parts of the utterance. For example, we could measure how long it takes someone to say a particular word or even a vowel within a word. In terms of the nature of the response, verbal responses are also a much richer source of data than button presses, which typically have at most seven different possible responses (and usually more like 2 or 3). With verbal responses, you can analyze acoustic information like intonation (as well as speech errors), which words the participant used (out of a potentially very large set of options), the syntax that the participant used, and so on. This task can be used to investigate processes that underlie speech production, of course, but it can also be useful for testing things that require a complex response in a more naturalistic way. For example, if you are interested in working memory and language, you might use a vocal response method to have participants verbally recall a word or sentence that they've recently encountered, rather than having them write it down (or select it from a set of possibilities).

TASKS THAT COLLECT DATA FROM (LARGELY) UNCONSCIOUS RESPONSES

Eye-Tracking

As the name suggests, eye-tracking involves tracking the movements of the eyes. To understand the usefulness of this kind of response, it's important to understand—at least in a general way—how the eyes work during normal scene perception and reading, so let's start there.

For the most part, our conscious experience of seeing things is that our eyes glide from thing to thing as we take in the world. When we want to look at something, our eyes move smoothly over to it, we gaze at it a moment or two, then move on to the next thing. However, this is not exactly how the eyes actually work. Instead, except when we are tracking an object that is moving through our field of view (like watching a car as it moves past us), our eyes take in the world in a series of jerky movements: they look (fixate) on one thing, then zip (saccade) to the next thing. This series of saccades and fixations is how we actually take in most visual information. Our eyes rest on something briefly, and then even more briefly, they move on.

In eye-tracking, cameras monitor the eyes and then computers calculate where the eyes fixate on a visual display, and for how long. This is a great method for reading because, unlike SPR, it requires no overt response (like pressing a button to reveal a new chunk of text) and, unlike other "online" measures like event-related brain potentials (ERPs) and fMRI, it can handle pretty naturalistic presentations of text. A researcher can, if he or she wants, just display a paragraph (or a single sentence) and record the pattern of fixations and saccades as someone reads it. Of course, this method is not completely naturalistic: participants in an eye-tracking study are likely to be wearing a headband that holds the cameras, or have to rest their head on a brace (or chin rest) in order to keep it steady. Some types of eye-trackers even require people to bite down on a "bite bar" to keep their head stable for cameras that are located on a table near the participant. Some eye-trackers don't require any of this, but participants must still keep their heads as still as possible. One example of an eye-tracking setup is shown in Figure 3.1. As you can see, it's unlikely that people forget that they're in an experiment.

In addition to allowing researchers to collect online data about language processing with language stimuli presented in a (relatively) natural way, eye-tracking provides a wealth of data, particularly compared with SPR. In SPR you generally get a few data points for each sentence, one for each time the participant presses the button to continue. But eye-tracking gives you a more complex set of data, which can be measured in a number of different ways. First, as with SPR, you can look at single words or sets of words, we can refer to these in general as interest areas. So, for example, you could measure the amount of time from when the participant first looks at an interest area until the participant

FIGURE 3.1 Head-mounted eye-tracker (EyeLink II, SR Research). Photo credit: Joshua McLawhorn.

moves on to fixate on another area; this is referred to as first pass reading time. You could also measure the length of the very first fixation in an interest area (this is generally used when the interest area is a single word); this is often called the first fixation time. These are fairly obvious types of things to measure and tell us something about how long it takes readers to read a word when they first encounter it, but we can also look at other things. For instance, we could look to see how long it took from when someone first looked at an interest area until they looked *ahead* into the text. This would include any regressive or backward-looking eye movements into previous text, and so would measure how long it took before a reader felt ready to continue on into the rest of the text—this measure is the regression path or go-past time. We could also measure the total amount of time someone spent fixating on an interest area before they indicated that they were finished with the whole trial, this is called (unsurprisingly) the total time. One could also measure the percentage of regressions out of an interest area, for example looking to see whether a certain aspect of a sentence made readers more likely to look back into previous text. Not every research study uses every measure, though these are a subset of the most commonly reported ones. In general, measures are thought to either reflect early processes during reading (like first pass) or early and later processes (like total time). Taken together, these types of measures can provide a fairly complete picture of how people take in information while they read.

While you can get a lot of important data just by placing text on a screen and analyzing eye movements, you can also be a little bit sneaky and create a dynamic display that changes based on where the participant is looking. This has been used to gain a better understanding of fundamental aspects of how information is taken in during reading (see Rayner, 1998, for an excellent overview). In one version, the boundary technique, one display is shown to a reader and then when their eyes cross an invisible boundary this display changes. This can be used to determine how important upcoming information is during reading. For example, the first display in this technique might be a sentence like "The waiter droqpcb the glasses," but when a participant's gaze crosses an invisible boundary at the end of *waiter*, then the display changes (during the saccade) to put the word "dropped" instead of droqpcb. Researchers can vary the words involved and how much is changed at the boundary to see what information influences whether readers take longer to read after the change.

Eye-tracking techniques are also not limited to studying written language, but can also be used to investigate spoken language. The Visual World Paradigm is a relatively new application of eye-tracking technology that has generated a lot of interest among researchers and has provided a new window onto how and when information is integrated during both word and sentence processing. In this technique, participants look at an array of pictures while listening to language stimuli and researchers may vary variables within the visual scene or the language. For example, one of the first studies to use this method examined the resolution of temporary referential ambiguities (Tanenhaus, Spivery-Knowlton, Eberhard, & Sedivy, 1995). For example, imagine that you hear "Put the apple on the towel in the box" at the same time that you see four real-world objects: a towel, an empty box, an apple that is on a towel, and a pencil. The pencil is not important, so we'll ignore it. But, the rest of the pictures matter: this sentence, combined with this set of objects, is at least temporarily ambiguous. When you hear "on the towel" it isn't clear whether you are listening to a description of the apple, or if you should put the apple on the other towel. Tanenhaus and his colleagues found that in this case, people look to the other towel quite a bit, compared with when there is a second apple (on a napkin) instead of that pencil. In this case, people look to both apples, settle on the apple on the towel, and then look to the box. In terms of when people look to objects (or pictures of objects), they do so very rapidly after identifying the word that labels it (if it is a noun). In the case of verbs, several studies have shown that people will make anticipatory looks to the most likely next word, based on knowledge about the verb and the objects in the picture. For example, Altmann and Kamide (1999) found that when hearing the word "eat" in a descriptive sentence like "The boy will eat the cake," people would look to the thing in the picture that could be eaten—even before hearing the word that labeled it.

These two applications of eye-tracking—reading and the visual world paradigm, mean that eye-tracking is well suited to studying both written and spoken language. In both cases it provides a rich data set of ongoing processing with relatively little intervention by the researcher. Participants are certainly aware of the eye-tracker, but the stimuli presented can be natural and no overt, conscious response from the participant is necessary (although many studies do still ask participants to answer questions about what they've read, to make sure they are paying attention).

Event-Related Brain Potentials

ERPs are a technique that allows us to match the electrical activity of the cerebral cortex to the presentation of stimuli to a participant. Like eye-tracking, it helps to understand this technique if you know a bit about the response we're measuring—in this case, the brain. Without going into a neuroanatomy lesson, let's cover a quick overview of the basic way in which the brain works. For our present purposes, the most important cells in the brain are the neurons, which use an electrical and chemical process to pass information around. The chemical part of this process you are probably already familiar with—it's this part that lets pharmacology companies develop drugs to help with things like depression: the drugs a patient takes go into the brain and alter the brain's "chemistry"—doing things like adding more of a certain type of chemical or making it harder for the brain to use excess amounts of another. (These chemicals are called neurotransmitters, and are siblings of hormones.) What we are interested in, though, is the electrical side.

Electrically, what's important is that the membrane of the neuron naturally maintains an electrical imbalance through the amount of positively and negatively charged ions in and near the membrane. It's actually the chemical part of the process that influences this imbalance—neurotransmitters influence how big the imbalance is. When the imbalance gets too big (too negative), that's when the electrical part happens: the negative imbalance triggers the neuron to send a wave of very, very, very small electrical change along the output extension of the neuron (called the axon). When this reaches the end of the axon (which can be almost as long as you are tall or can be pretty much nonexistent), the electrical change triggers a release of neurotransmitters, which in turn changes the electrical imbalance of the next neurons in the chain.

This is a pretty simplified view of the process, but it helps to have a sense of what exactly ERPs are recording: when populations of neurons in the cortex are physically aligned in parallel and become electrically active at the same time, this creates a small electromagnetic field that can be recorded by an electroencephalograph and gets visually represented as brain waves. For example, sleep researchers use electroencephalograms (EEGs) to help understand brain activity while we sleep. ERP researchers record EEGs, but then also digitize them and time lock them to external events—for example, they might mark the point in the EEG when someone saw the word "dog."

Here's the extremely cool thing about the brain—despite all our different experiences in life and with language, our brains respond in remarkably similar ways when confronted with the same linguistic information. For example, in some pioneering work in the late 1970s and early 1980s on using ERPs to study language processing, Marta Kutas and Steve Hillyard were interested in whether surprising language would elicit the same response as other surprising visual stimuli. Previous work in nonlanguage research had found that unexpected stimuli elicited an increased positive polarity around 300 ms after the onset of the stimulus. They wanted to know whether unexpected linguistic stimuli would elicit the same response, or whether it would be different. They presented people with sets of sentences that sometimes contained words that were sensible, but unexpectedly presented in all caps, or words that were semantically nonsensical but in the regular font, or both of these things. For example, people would see sentences like "The warm toast was spread with socks" or "The warm toast was spread with BUTTER." They found that when words were unexpectedly in uppercase, then an ERP response similar to that found for other unexpected stimuli was found. However, semantically odd words gave a different, new response: a sharp, negative spike in electrical activity starting around 350 ms after the anomalous word (e.g., "socks"). This response peaked around 400 ms and then quickly subsided. This is called the N400 component (for negative, with a peak at 400 ms). In numerous other studies, it has been discovered that all words elicit an N400 response, but the size of this response depends on things like how expected a particular word is and how much the actual word deviates from the expected word. The N400 response has come to be thought of as an index of semantic and pragmatic processing (see Kutas, van Petten, & Kluender, 2006; for a review of recent findings from ERP studies on language). The N400 is interesting for a number of reasons, but for current purposes, this example shows us how ERPs can be used to study language processing—from these results we know that people begin to process the meaning of the words they encounter by at least 350 ms after they first encounter them. We also know, from many other studies, that the processes that support this lexical processing appear to be distinct from (though possibly interacting with) the processes that support processing the structure of sentences. We know this because when you present people with syntactically strange or ungrammatical sentences, you generally do not see an increased N400 response (although there are a few studies that have found them) but you do find a different response instead. Examining ERP responses allows researchers to both uncover the time course of language processing, and to tease apart different processes that underlie it.

While ERPs are a powerful tool for investigating language processing, they do have some aspects that make them less than completely perfect. First, they cannot tell us where in the brain a signal originated. This is called the inverse problem: we don't know how many different populations of neurons are involved in producing the final pattern of electrical activity that we see at the scalp. Because of the nature of how electrical sources interact, any given scalp pattern could be generated by any number of possible activations in the brain itself. Hence, the problem. This problem has been mitigated somewhat in recent years by the use of dense electrode arrays and some very fancy (and often expensive) mathematical algorithms that can give the most probable locations of the source of the scalp pattern (called source generators). Because of other neurological research using other techniques (e.g., fMRI), we know a bit about which parts of the brain are most likely to be involved in language processing and which are not. This helps narrow down the possibilities for the location of the source generators for ERPs. However, in the end, this problem will remain for ERPs and their natural strength is in discovering the time course of language processes rather than in locating where in the brain the processes originate.

Another issue with ERPs is that the signal that the brain produces is very, very small. Some noise is introduced when that signal is amplified. Further noise comes from external electrical noise (we are all bathed in electromagnetic waves from various sources like electrical outlets, computer monitors, even the elevator down the hall). Even more noise can be introduced from non-cognitive electrical activity in the body: when a muscle moves, that produces an electrical signal; the eyes are a particular issue here because they are physically close to what we want to measure and can overwhelm the signal when we move them. Finally, there is "cognitive" noise as well—the brain is doing more that just processing language, and even the best, most focused participants are going to occasionally get distracted or lose focus over the course of the hour or more they spend in the experiment. So, all of these sources of noise make recording and analyzing ERP data trickier than it might seem at first. To help improve the signal-to-noise ratio, ERP studies of language typically have anywhere from 30 to 50 experimental items per condition. This means that a study with four conditions could have as many as 200 items that each participant must look at. Even though participants are typically seated in a comfortable chair and encouraged to relax, the sheer number of trials can be somewhat tiring. Nonetheless, as with eye-tracking, ERPs have provided researchers with a wealth of important data and will certainly continue to do so.

Functional Magnetic Resonance Imaging

This technique provides researchers with data about which areas of the brain become active during one task (say, reading a verb) compared with another task (typically a "control" task that establishes a baseline level of activation, like looking at random strings of letters or letter-like symbols). Activation in this case is measured by changes in blood flow. In many ways, fMRI can be considered the complement to ERPs. While ERPs have the potential for accuracy on the order of milliseconds (and better) and so can provide an excellent source of time course information, fMRI resolution is on the order of millimeters and so provides an excellent source of localizing the areas in the brain that underlie language processing. As with ERPs, a full discussion of the technique is beyond the scope of this chapter, but, basically, fMRI works by taking advantage of the magnetic properties of the hydrogen nuclei that are present in the water molecules in our body, including the brain.

Similar to a bar magnet, hydrogen nuclei are sensitive to magnetic fields due to their very small positive charge, and will align themselves with a strong magnetic field. You may already be familiar with MRI, which gives doctors excellent, three-dimensional (3-D) views of the inside of the body, including soft tissue. In MRI, a very powerful electromagnet aligns the nuclei and a carefully tuned, brief radio pulse disrupts this alignment, causing the nuclei to rotate up to 180° out of their alignment with the magnetic field (this is called excitation). Researchers can fine-tune the radio pulse to cause the alignment to be big or small and then can measure the time it takes for the nuclei to realign or "relax" back into the steady magnetic field. Crucially, different types of body tissue have somewhat different magnetic sensitivities and this means that the hydrogen molecules will relax in different tissue types at different rates. Doctors and researchers take advantage of these differences to build 3-D representations that distinguish different parts of the body. In the case of the brain, we can thus distinguish between the skull, cerebrospinal fluid, gray matter, and white matter. This gives us structural or anatomical MRI—we can map out the brain (or other body parts) to the square millimeter.

So, what about *functional* MRI? It works largely the same way, except that researchers measure what's called the BOLD response—blood oxygen level-dependent signal. Simplifying a bit, the BOLD response takes advantage of two things: first, iron molecules carrying oxygen in the blood have a different magnetic response than iron that is not carrying oxygen—this allows us to distinguish oxygenated blood from deoxygenated blood. Second, groups of neurons becoming active at the same time together tend to use more oxygen than neurons that are at rest, and so about 2 s after an increase in neural activity, there is a large increase in the amount of oxygenated blood that goes to those neurons (in fact, more oxygen is usually sent than is strictly needed). BOLD detects a characteristic spike in increased blood flow that peaks sharply around 6 s after the increase in neuronal activity.

So, we can measure changes in blood flow in the brain in response to changes in neuronal activity. Researchers map this onto anatomical MRI, and voila—we can see which parts of the brain become more active in response to, well, whatever stimuli we want to present. But, because blood flow changes are relatively slow, we can't distinguish events in time quite as well. This is how fMRI is a complement to ERPs—while ERPs can give us information about when electrical activity changes in the brain, fMRI can tell us where it happens. Often, it is not simply one place in the brain that shows changes in activity level. For example, a couple of studies (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Menenti, Petersson, Scheeringa, & Hagoort, 2008) have shown increases in activity in a region of the brain called the left inferior frontal gyrus (a region in the frontal lobe of the left hemisphere that includes Broca's area) when sentences are presented, which contains information that violate our knowledge of the real world (e.g., *With the lights on you can see less at night* compared with *With the lights on you can see more at night*). However, increases in activation in these violations of world knowledge are not only found in the left hemisphere, but also in the right hemisphere counterpart (right inferior frontal gyrus), suggesting that both hemispheres play a role in evaluating sentences with respect to real-world knowledge.

Designing Experiments

These techniques make it possible to know a lot about how language is represented and processed, but by themselves they are not useful—they are only part of a larger experimental design that produces results. Each of the kinds of response we've covered above has special considerations for experiments using that response, but what we'll do in this section is have a brief discussion of the general considerations that are taken into account by language researchers when designing experiments. Many of these are also true more generally of any cognitive psychology experiment, but some are more specific to having language as your object of study.

Strength in Numbers

While researchers use one or two examples to show what their stimuli look like, they use a lot more than just one or two items in the actual experiment. This is because researchers need to have data from many responses to the same kind of stimulus in order to be sure that the response they get isn't just due to: (1) chance, (2) unintended idiosyncratic associations, (3) unintended aspects of the stimulus, and (4) other external factors. For example, let's say we want to know if people are able to process color names faster when they refer to basic (primary, secondary) colors compared with nonbasic (tertiary and beyond) colors. So, will people respond faster to "red" compared with "magenta"? Let's say we test how quickly someone can say the words—a verbal response. With only a single person and a single set of test items, then the following problems can arise: (1) someone might randomly take longer to respond, (2) maybe red is someone's favorite color, (3) "red" is shorter than "magenta," and (4) someone might need to sneeze right as they are supposed to respond. These problems can be alleviated somewhat by running the same item set on lots and lots of people, and there are a few studies out there that have done just this. However, the solution that works much better for the vast majority of experiments is to create multiple sets of items that are representative of the type of factor you want to test. Sticking with our color example, we would want to test several primary colors and several nonbasic colors, not just one example of each color type.

How many sets of items does an experimenter need? This is actually more complicated than it sounds. The number of items needed to ensure that you have really gotten a big enough sample depends on the technique used and the design of the study. As mentioned earlier, ERP studies typically use between 30 and 50 items for each condition that is tested. So, let's say you were interested in how the language processing system responds when it encounters something nonsensical. You could have a study with two conditions: sentences that are sensible and sentences that are not. Ideally, these sentences would differ as little as possible in every other way (see the next section) and so you might have a sentence set like: *The boy rode his bike down the street* and *The boy rode his bike down the tree*. You would also need to create and test at least 59 more sentence sets that differed in the same way. Eye-tracking studies typically need around 10 items per condition and so the same study with this different method would only need 20 sentence sets.

AVOIDING CONFOUNDS: CONTROLLING FOR KNOWN PROBLEMS AND TRYING TO AVOID UNKNOWN ONES

Another important consideration that researchers must take into account is the potential influence of unwanted factors. The easiest way to see this is when you are comparing two groups of people. Let's say you are interested in comparing how nonsensical sentences are processed by native speakers of a language compared with advanced second learners of that language. You would want to be careful to test two groups of people that are otherwise as similar as possible, because otherwise you would run the risk of having some other difference between the groups contribute to any difference in processing. For example, let's say that you don't control for the ages of your participants and that because of the way that you recruited your participants (or even by bad luck) you have a group of native speakers who are all between the ages of 18 and 23, but your second language learners are between the ages of 30 and 35. If there are differences between the groups, you cannot be sure whether it is due to their language background or their age. There are some statistical tricks you can do to deal with this, but it is easier (and a better design) to just make sure that age is held equal between the two groups to begin with.

These same concerns apply to materials in psycholinguistic studies as well—and in an analogous way. We are sampling a subset of individual people from a larger population, and we are also sampling individual language items from a large population (of language). It turns out that there are a number of properties of language that could, if ignored, cause problems for correctly determining the outcome of an experiment. Sticking with our nonsensical sentences a bit longer, imagine that you dutifully create a set of 20 sentences for an eye-tracking study. However, you focus so much on making sure that the word change in each sentence set results in something nonsensical that you don't notice another pattern: all the critical, target words in the nonsensical condition are longer than the words in the sensical versions of the sentences. Oops. It turns out that word length makes a contribution to certain measures of reading time. Other known factors include the frequency with which the word occurs in language use and the semantic relatedness between words. So, unless you are specifically interested in testing (and manipulating), say, word length or frequency, you would want to make sure that your critical words did not differ in a systematic way between conditions with respect to these factors.

This makes the design and development of linguistic items somewhat challenging, and differences in opinion between researchers about the validity of each others' materials has occasionally cropped up. Sometimes genuine mistakes are made, but other times a research team may simply not take into account some potential difference that turns out to be important. Another team spots the difference and tests the hypothesis that it was this previously unknown factor that caused or contributed to the outcome.

Being Sneaky: Filler Items

Another aspect of psycholinguistic experiments that is important, but not generally given too much airtime in descriptions of results, is the use of additional items that are not a part of the experimental design, but serve to distract participants from the true nature of the design or help ensure that they give valid responses. For example, let's say that you were interested in how quickly people are able to identify words and what factors are implicated in that process. You might well use a lexical decision task, in which people are asked to decide whether a string of letters is a real word or not. They may press a button for "yes" and another for "no" and the timing of their correct responses will be analyzed. Even if you are not interested in how people respond to nonwords, it is important to include some nonword trials (in which a participant might see the “thripe” or “malk”). To appreciate why, imagine you were a participant in the study. If all the letter strings you see are real worlds, you might well stop actually doing the task and just start pressing the “yes” button. Now your data no longer reflects the process of deciding whether something is a word—it reflects how fast you can push the “yes” button. In other cases, filler items are important to obscure the nature of what the researcher is interested in. Let’s say that now you’re interested in how much a particular kind of grammatical error disrupts processing. Because you want your experimental items to be as similar as possible to avoid unintended effects, you have a set of materials in which the error, when it occurs, is always in the same place in the sentence. It won’t take long for participants to anticipate that they might see an error in this spot, and so you may find that they start reading differently in anticipation of potential errors. This is a problem because you want people to read as normally as possible—after all, you are interested in how disruptive the error is to normal reading processes, not how disruptive it is to “anticipating an error any moment” reading processes. So, you will probably need a lot of filler (it is not unusual to see experiments with twice as many filler items as experimental items), and you will probably need to have some of those filler have errors in other parts of the sentence. By varying the filler carefully, you can obscure the specific purpose of your experiment and keep participants from guessing what’s coming next. As a bonus, this also increases the likelihood that they will stay interested and alert while reading your sentences.

Balancing the Numbers

In the “Strength in Numbers” section earlier, we covered the fact that researchers need to present multiple items of the same language phenomena. Crucially, they generally don’t present the same item in different versions to the same participant (although it has been done on occasion, and certain types of methods, for example, those used with speech perception, can do this without a problem). Instead, they will often divide their sentences in groups and have participants see a subset of their materials. For example, say you were still interested in the effects of grammatical errors. You make 20 sentence pairs, one of which might look like this:

Grammatical: The key to the cabinets is on the desk.
Ungrammatical: The key to the cabinets are on the desk.

If you show both of these to the same participant, they will undoubtedly notice and start to get suspicious about why they are seeing two pretty much identical sentences, particularly if you show all 40 of the sentences (20 in each version) so that every sentence that they see has a paired sentence that they see later. So, you show the grammatical version of this sentence pair to half of the participants and the ungrammatical version to the other half. Between these two groups, you present half of the sentences in their grammatical version and other half in their ungrammatical version. This is sometimes referred to as a repeated measures design and is certainly not unique to psycholinguistics, although it is used quite often.

Even when this particular way of balancing the numbers is not used, researchers are careful to ensure that the way they present their materials to their participants doesn't result in an unintentionally unbalanced design, or one that inadvertently causes some type of unwanted strategy on the part of the participants.

REFERENCES

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247–264.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*, 438–441.

Kutas, M., van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II (1994–2005). In M. Gernsbacher & M. Traxler (Eds.), *Handbook of psycholinguistics* (pp. 659–724). New York: Elsevier Press.

Menenti, L., Petersson, K. M., Scheeringa, R., & Hagoort, P. (2008). When elephants fly: Differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *Journal of Cognitive Neuroscience*, *21*(12), 2358–2368.

Mitchell, D. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (pp. 15–32). Brighton, UK: Psychology Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Tanenhaus, M. K., Spivery-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634.

4

Information Flow and Language Ambiguity

LANGUAGE IS OFTEN AMBIGUOUS

Despite the obvious desire for language to be clear and easy to understand, language is often ambiguous. In fact, it is so often ambiguous that this bears repeating: Language is often ambiguous. What does this mean? Take the following sentence:

1. I went down to the bank yesterday.

So, where did I go yesterday? Did I go to a place where money changes hands, or did I go to a place where water flows nearby? Now, you might (correctly) object that this sentence is unfairly out of context and that in any real world situation the hearer would know exactly which "bank" the speaker meant. Certainly, context can and does help in the final interpretation of the meanings of words, but researchers have been especially interested in the processing of language as it unfolds over time, and so while context is clearly involved in coming to the correct final interpretation of a word (or sentence), a major question concerning information flow has been about what happens in those first moments when you encounter a word. In particular, which meanings immediately become activated? Context might help in this case, but the key question is, when? And what about when there isn't any context? What happens if I come up and just say "bank"?

Ambiguity isn't limited to single words, either. As we saw in chapter 2, sentences have structure beyond the linear order of the words in them. And, it turns out that sometimes the same string of words can have more one possible structure. Take the following:

2. I watched the man with the binoculars.

So, according to this sentence, who had the binoculars? Me? The man? Both are possible. What researchers have wanted to know is how people initially interpret this sentence when there isn't any context and whether this interpretation changes when there is context.

Finally, the thing that really makes sentences particularly ambiguous is the fact that sentences can have more than one possible syntactic structure temporarily, for some brief period of time during the sentence itself, but that by the end of the sentence only one structure is possible. For example:

3. The florist sent the flowers was very pleased.

Here, the sentence has one and only one syntactic structure by the end—in which *the florist* is the subject and *was very pleased* is the predicate. In the middle, we have *sent the flowers*, which is a clause that gives more detail about the florist (it is a reduced form of "who was sent the flowers"). However, there is a moment in the sentence when this final structure is not the only possible one—for a moment there, *sent the flowers* could be the predicate of the sentence, not just a clause that gives more detail. The question this kind of ambiguity poses is: Does the language processing system notice the temporary ambiguity? If it does, how does the system handle it?

WHAT INFORMATION DO WE HAVE ACCESS TO AND WHEN?

Psycholinguists have spent a great deal of time and research energy on these types of ambiguities in language, but this is not because they are interested in ambiguities alone. The big question at the heart of this chapter, and in the research on ambiguity, is what kinds of information do people use to process language, and whether (and how) these different sources of information interact. This is both a current topic in psycholinguistics and an "old" topic in the sense that efforts to answer this question have driven some of the seminal work of the field.

The Need for Speed Revisited

In chapter 1, we discussed the need for speed in language processing. The speed required by language in fact creates a tradeoff between accuracy and speed—we need a system that is both fast and accurate, but speed may come at the cost of accuracy. Over the years, there have been two main hypotheses about how to deal with this speed/accuracy tradeoff. On one side, the hypothesis has been that the best way to be as fast and accurate as possible is to have a system in which there are clear, limited information paths with multiple, separate processes that are specialized to handle each part of the language process individually. This is the modular approach. On the other side, the hypothesis is that the system gains speed and accuracy by allowing multiple sources of information to interact as they are available—this is the interactionist approach.

Orthogonal to these two approaches, but interacting with them and just as important, is the way in which information flows—does it only flow in one direction? Or can information from "higher" processes feed back to influence "lower" processes, and/or can information flow through multiple processes concurrently? When the processing system allows information to flow in only one direction from one process to another, it is said to be a serial system. When the system allows multiple processes to co-occur in time and for information from later processes to flow back and influence previous (but still ongoing) processes, then it is called a parallel system.

In this chapter, we will discuss the modular and interactionist approaches in detail as well as examine some of the experimental evidence in favor of each. To do this, we will focus on an area that has been at the center of the debate between these approaches: processing ambiguous words and sentences. Let's start with words.

BUGGING BUGS TO LOOK FOR BUGS: HOW WE HANDLE MULTIPLE WORD MEANINGS AND CONTEXT

Before we can turn to ambiguous words, we first need to talk a little about how people recognize and understand words when they are not ambiguous. This is a huge area of interest and a lot of factors appear to influence this process of recognition and meaning. A summary of four major factors is given in Table 4.1.

Another important thing to know about what happens when we encounter a word is that we cannot help but process it—it is an "automatic" process. This is true for written language as well, for people who can read. A very famous example of this is the Stroop Test—in which people are given a list of words printed in different colors. They are asked to ignore the words

TABLE 4.1 MAJOR FACTORS THAT INFLUENCE HOW QUICKLY WORDS ARE ACCESSED

Factor	Description	Effect
Frequency	The number of times someone has encountered a word	More frequently encountered words are responded to more quickly
Age of acquisition	The age at which someone first learns a word	Earlier learned words are responded to more quickly (but tricky because often confounded with cumulative frequency)
Neighborhood size	Number of other words that are similar in form (e.g., *lime* and *time* are neighbors)	The more neighbors a word has, the more quickly it is responded to
Semantic priming	Encountering a semantically related word immediately before a word (e.g., seeing *doctor* right before seeing *nurse*)	Seeing a related word speeds response times

themselves and only name the colors that the words are printed in. It seems easy enough, but when the words themselves are color terms, so that the word "red" is printed using blue ink or the word "yellow" is printed using the color red, people have some difficulty—as shown by longer times to name the color and even errors—starting to say the word itself instead of the color. You can try this yourself. The Stroop Effect occurs because reading is automatic, but determining and naming colors is not. So, these two processes interfere with each other, and the identity of the words can affect how easy it is to say the color name.

ON TO AMBIGUOUS WORDS

So, what about when the word that you are reading automatically has more than one meaning?

There are a few logical possibilities: Prior context could have no effect on initial access—this is the autonomous access model. The benefit of this model is that it can be fast. Or, prior context could constrain access in a way that allows only the context-appropriate meaning to be activated—this is the context-sensitive model. The logic for this model is why waste energy accessing more than what you need given the context? There is also a third possibility, a combination of the first two: this is the reordered access model, in which prior context has a limited effect: it increases the accessibility of the appropriate meaning, but doesn't narrow down access to only that meaning. So, all meanings are accessed, but the context-appropriate one is accessed earlier.

Initial investigation into this issue found evidence in favor of autonomous access in which context did not prevent alternative meanings from becoming activated. For example, Onifer and Swinney (1981) used a cross-modal method in which participants listened to sentences that provided a biasing context toward one meaning of an ambiguous word or the other, and then, at that ambiguous word, they saw a different word on a computer screen that they needed to make a lexical decision to. Onifer and Swinney found that words that were related to *either* meaning of the target word prompted faster response times than words that were semantically unrelated, regardless of the context that the target word was in. For example, participants would hear a sentence like "The postal clerk put the package on a postal **scale** to see if it had enough postage," and then as they heard the target word *scale*, participants would see one of four possible probe words on the screen in front of them: two words related to the different meanings of the target word (*fish* and *weight*) and unrelated comparison words (*coal* and *source*). They would then have to respond as quickly as possible to indicate whether they thought the probe word was a real word of English or not. In critical trials, the probe word was always a real word and so what we are interested in is how quickly participants responded. Crucially, they were faster to respond to *fish* (compared to *coal)*, even though the ambiguous word *scale* in this sentence very clearly refers to the kind of scale that weighs things and not something found on a fish. Further, the size of the increase in response speed for *fish* was basically the same as the increase in response speed for *weight* (compared to *source*). The fact that both *fish* and *weight* cause faster times supports the idea that in terms of information flow, sentence-level information is not initially used during the process of accessing a word. If it were, then we would expect to see benefits for context-appropriate meaning at the word itself—*weight* should show a larger increase in response speed than *fish*. These results are thus consistent with the autonomous access model, sometimes also called exhaustive access because in this model all meanings of a word are accessed.

Interestingly, an important factor for ambiguity resolution appears to be the frequency of the different meanings of the ambiguous words. Some words, like "yarn" have a frequently used meaning (something you knit or crochet with) and infrequent meaning (a tall tale) while others have meanings that are all roughly equal in frequency. We can distinguish these two types of words by calling them balanced (both meanings roughly equal) and polarized (one meaning more frequent than the other(s)). While Onifer and Swinney did look at polarized words, they didn't find any difference between the two meanings. However, later experiments have found that meaning frequency is important, and many studies have found a *subordinate-bias effect*. This effect is as follows: in a neutral, non-biasing context, words that are balanced cause longer reading times than words that are either unbalanced or unambiguous. But, if later context disambiguates in favor of the less frequent ("subordinate") meaning, then reading time increases at that point. So, this suggests that in an initially neutral context, only the most frequent ("dominant") meaning is accessed for a polarized word. What about in a context that biases toward the dominant or subordinate meaning *before* the word? In this case, there are longer reading times for the unbalanced word when the context is biased toward the subordinate meaning (compared to context biased toward the dominant meaning or an unambiguous word). (See Sereno, O'Donnell, & Rayner, 2006, for a larger description.) This subordinate-bias effect has been studied quite a bit in order to determine which model best accounts for it and effects related to it. Importantly, in terms of our models of meaning activation—a reordered access model would account for the subordinate bias effect by stating that the subordinate meaning may, with context, become as accessible as the dominant meaning, but the dominant meaning will still become activated, even when the context is strongly biased toward the subordinate meaning. This means that frequency cannot be entirely overruled by context.

Some natural questions to ask at this point are: how do we know what a dominant meaning is, and how biased is a "biasing" context? In fact, both of these questions have posed problems for researchers: first, meaning dominance is usually determined by questionnaire studies that ask people to judge the meanings of words in various ways. Over the years there has been disagreement among researchers about materials, and part of the problem is that word use can be influenced by local conditions. For example, Rayner, Binder, and Duffy (1999) point out that while in Florida the most frequent use of the word "screen" is reported to be referring to screen doors, in Massachusetts the most frequent use is to refer to movie screens.

Biasing contexts are even more tricky. Martin, Vu, Kellas, and Metcalf (1999) provided data in which strongly biased contexts cause the subordinate bias effect to go away, supporting the possibility of a selective model in which context may have a bigger role. These authors argue in favor of a context-sensitive model of ambiguity resolution in which the strength of the bias in the context can overcome frequency bias, allowing context to have a larger influence when it is sufficiently biasing. To test whether this is true, they had participants do a self-paced reading task with short two-sentence passages as in the following:

4. The custodian fixed the problem. She inserted the **bulb** into the empty socket.
5. The gardener dug a hole. She inserted the **bulb** carefully into the soil.

These are examples of strongly biased contexts, with the ambiguous (target) word in bold. They compared these to weakly biasing contexts like the ones in the following:

6. The farmer saw the entrance. He reported the **mine** to the survey crew.
7. The scout patrolled the area. He reported the **mine** to the commanding officer.

In one experiment, they added a second, naming task. In this case, a related or unrelated word would appear immediately after the participant had pressed the button after reading the target, ambiguous word. For example, after reading "bulb" in the first sentence, participants might see "light" or "tavern." Martin and his colleagues found that the weakly biased contexts replicated the subordinate bias effect, but that this effect went away in strongly biased contexts. This is evidence in favor of the idea that context does play a role in accessing the meaning of words.

More recently, Sereno et al. (2006) have questioned the words that previous researchers have used as semantically unrelated control words. It turns out that in the vast majority of studies, response times to related words are compared to response times to these unrelated control words. Also in most studies, each control word is matched in frequency to its corresponding ambiguous word. Remember that frequency plays an important role in processing words, with high frequency words processed more quickly, all else being equal. So, it is very important that the control word and the ambiguous target word be matched in frequency. However, the problem (as Sereno et al. see it) is that the control word has been matched to the *overall* frequency of the ambiguous word, counting all uses of that word regardless of specific meaning. So, for example, if one meaning of an ambiguous word was used 50 times in a set of a million words, and the other meaning was only used 10 times, then a control word with a frequency of about 60 would be selected, because the total frequency of the ambiguous word's use is 60. Sereno and her colleagues instead compared ambiguous words to control words that matched the frequency of the ambiguous word's individual meanings. So, Sereno et al. compared two control words, one with a frequency that was the same as the subordinate meaning (e.g., 10) and one that matched the overall word use (e.g., 60). They found that this made a difference. They used an eye-tracking methodology to examine reading times for sentences in which the context was biased toward the subordinate meaning of a polarized word. First, they found a subordinate-bias effect for the target ambiguous word to the overall-frequency control word—meaning that the ambiguous word showed longer reading times than the control. This replicates and confirms previous work. However, when they compared the ambiguous word to a control word that was matched to the frequency of the subordinate meaning, then the opposite was true—the ambiguous word showed *faster* reading times than the control word. It is not yet clear why this reverse effect happens, but this result highlights both the importance of minute details in experimental design and why we are still not certain to what extent context exerts an initial effect on meaning retrieval in word processing.

Bottom-Line on Ambiguous Words: Context Counts

One thing that experimental evidence does converge on is that context does influence how meanings of ambiguous words are accessed. There is a still an ongoing debate about the extent of this influence, and in particular whether it can override frequency information sufficiently to suppress highly frequent meanings. But, for our purposes here, the key, take-home message at this point is that despite early evidence against the role of sentence context in word meaning, it does appear that context can, at minimum, modulate the availability of appropriate word meanings.

GARDEN PATHS AND RACING HORSES: HOW WE HANDLE MULTIPLE SENTENCE STRUCTURES

The same issues that we saw with respect to ambiguous words play out in similar ways at the level of entire sentences. This time the question is—when and how does context influence how sentences are constructed (instead of how meanings are accessed)? Every word can have a grammatical category assigned to it-noun, verb, adjective, and so on. As we discussed in chapter 1, different languages impose different rules about how grammatical categories may be combined. These syntactic rules generally apply without too much consideration for how sensible the meaning of a sentence is. So, we can have a sentence that is grammatically well-formed, but meaningless:

8. Swift tables never fly below three dogs in the afternoon.

And this is importantly different from a string of words that conveys a meaning, but is not grammatically well-formed:

9. Use to car need we go store to.

So, when people are processing a sentence, they need to take words and assemble them according to the rules of the language. We are reasonably sure, based on many experimental results, that people process sentences in a largely incremental fashion. This means that as we encounter each word we try to process it as fully as possible. From a working memory perspective, this makes a lot of sense. If we can structure words as we encounter them, then we are effectively "chunking" them into bigger units, which makes them easier to keep track of. However, this incremental approach means that we run the risk of being wrong more often than if we took a more "wait-and-see" approach to incoming words. Upcoming, future words might provide disambiguating information that is inconsistent with how we initially build the sentence. Nonetheless, we seem to take the risk and interpret words and incorporate them into structure as we go. Key questions about how we do this include: Can context, particularly semantic and pragmatic information, influence how this happens? As with ambiguous words, ambiguous sentences give us an effective tool to investigate this. If there are two (or more!) possible structures to build, what happens? What kinds of information influence structure building, and when?

Before discussing the two main models that address these questions, let's return to a more full description of the two types of ambiguities. First, there are standing ambiguities: this is the case when there is no necessary reason to rule out or pick a given structure, and so the sentence has two or more possible structures even at the final word.

10. Sam watched the spy with the binoculars.

Here, it is uncertain from this sentence alone who had the binoculars—Sam or the spy. In terms of structure, this sentence is ambiguous because the preposition phrase *with the binoculars* can be attached either as a phrase that modifies *the spy* or attached to the verb phrase itself, in which case it modifies how the watching was done. Simplified syntactic structures for these two interpretations are given in Figure 4.1.

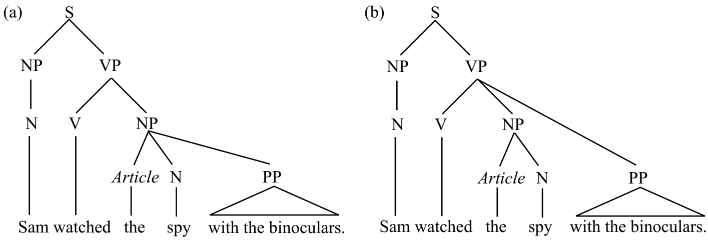

FIGURE 4.1 Two alternate structures for *Sam watched the spy with the binoculars*. (a) Shows the structure for the interpretation in which the spy has the binoculars. (b) Shows the structure for the interpretation in which Sam has the binoculars.

Next, let's return to temporary ambiguities. These are harder to detect because they are, as the name suggests, temporary. They occur when two or more possible structures are possible in a sentence but that at some point one structure is unambiguously the correct one. This point is called the disambiguation point. Take for example:

11. The horse raced past the barn fell.

This is the canonical example sentence for temporary ambiguities. Upon first encountering this sentence, many people find it difficult to get the correct interpretation of it, or even believe it is ungrammatical. So, before dissecting this sentence, let's look at another sentence that is syntactically the same:

12. The glass submerged in the water cracked.

Here, the interpretation of the sentence is a bit clearer—there is a glass that was submerged in water and this same glass cracked. We can now apply this to the horse in (11): there is a horse that was raced past a barn, and this same horse fell down. Does context matter in how we process these sentences? The very fact that (12) is easier than (11) is an important clue.

But, how are these sentences ambiguous? They are both examples of *reduced relative clauses* and depend on English's ability to omit certain words when forming a relative clause. A full version of these sentences would look like this:

13. The horse that was raced past the barn fell.
14. The glass that was submerged in the water cracked.

Crucially, both the *that* and the *was* can be omitted in English—this creates the ambiguity by making the set of words for the relative clause temporarily indistinguishable from a sentence without a relative clause, at least for many verbs. Example:

15. The horse raced past the barn.

We can end this sentence after *barn* because up to this point there is another syntactic structure that can be applied, in which *raced* is the main verb of the sentence, not part of a description of the horse, and so the horse is doing the racing, and then an additional phrase tells us where. In fact, we could even potentially end the sentence after *raced*.

So, back to the main questions: What do we do when we encounter these ambiguities? Do we entertain all possible interpretations or do we pick one and stick with it until forced to abandon it? Can prior context help disambiguate, and if so—when does it have this effect? These questions are remarkably similar to those posed for ambiguous word meanings, and the answers that have been proposed are also quite similar. In the case of structural ambiguity, there are two main models that have proposed: The garden path model and the constraint-based model. The garden path model is similar to the autonomous access model for word meaning—in this model prior semantic context has no initial effect. The constraint-based model is similar to the context-sensitive model in word meaning—in this model all kinds of information, including semantic context, can bias initial interpretation. Another big difference between these two models is that the garden-path model is a two-stage model while the constraint-based model is a one-stage model. This means that in the garden-path model processing happens in two distinct steps while in the constraint-based model everything happens as part of the same process.

Being Led Down the Garden Path . . .

As with the exhaustive access model for word processing, the garden path model achieves speed by simplifying the process: Context does not have an immediate impact on building sentence structure. Temporary ambiguities like the one in (11) are often referred to as garden path sentences, as in the comprehender is being led down (and abandoned in) the garden path to the wrong place. In the garden path model (e.g., Frazier, 1987; Frazier & Rayner, 1982), sentence processing happens in two stages: an initial structure building stage in which the only information that is used is syntactic, and then a second stage in which the structure is checked against semantic and pragmatic information. Importantly, in the first stage, only one structure is constructed, even if others are possible. Then, if the structure is subsequently incompatible with further syntactic information, or semantic or pragmatic information, the structure is revised in the second stage.

The crucial part of this model, and what really distinguishes it from other models, is this first stage. If the words coming in came from more than one structure, but the processing system (or parser) only builds one structure, a good question is what structure does the parser build, particularly because according to the model, no semantic or pragmatic information is available at this point during processing. The garden path model has two key principles that apply to how words are initially attached to the sentence structure: minimal attachment and late closure. Let's start with minimal attachment first.

Very simply put, minimal attachment says "keep it simple." This means that incoming words should be attached to the current sentence structure using as few "nodes" as possible. This is an excellent time to remember that linguists don't agree yet on the best way to characterize syntactic structures, and so this part of the model is somewhat tricky if one tries to apply current syntactic theory to it. But, using a simplified version of syntactic structure building, we can see how this principle can apply during processing. Let's take the beginning of the following sentence shown with two possible structures in Figures 4.2 and 4.3.

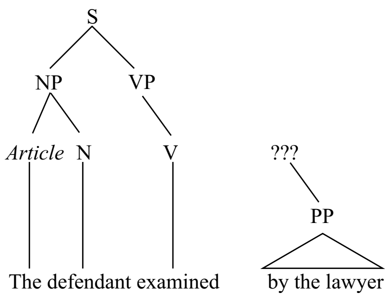

FIGURE 4.2 Initial structure build for *The defendant examined*, leaving no clear place to attach *by the lawyer*. In this case, a different structure must be used.

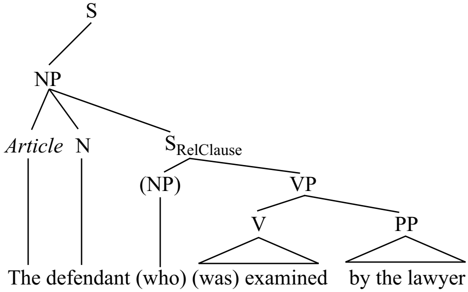

FIGURE 4.3 Another structure for *The defendant examined* that is more complex but allows *by the lawyer* to be attached.

The defendant examined by the lawyer...

Figure 4.2 shows a simple structure in which *the defendant* is assigned to the subject position at the first NP and *examined* is assigned to be the main verb of the sentence. The problem is that when readers encounter the beginning of the PP, *by*, it is unclear what to do with this phrase. Figure 4.3 shows a more complex structure in which the verb *examined* is assigned to be the verb in a relative clause that modifies *the defendant*. Notice here that in order for this interpretation to be correct, we must basically posit some unspoken words that are shown in parentheses, *who* and *was*, and build a bunch of extra structure to accommodate the relative clause. The key idea here is that the second structure in Figure 4.2 is more complex than the first, and so not built by the parser initially because it does not follow the strategy of minimal attachment. Unfortunately for the comprehender in this case, the first structure, while simple and following minimal attachment, is not ultimately correct and so according to the garden path model, this initial structure is recognized as incorrect at the word *by*, and the sentence structure must be revised at this point to match the actual input.

The second principle of the garden path model is most clearly illustrated by standing ambiguities. The principle of late closure basically states that incoming words should be attached to the current phrase if possible—in other words, the parser should "attach low" when trying to build incoming words into the current structure. This explains why there is a preference in English to interpret the binoculars as being with the spy in the sentence (10), repeated as (16), rather than Sam.

16. Sam watched the spy with the binoculars.

Recall from Figure 3.1 that *with the binoculars* may be attached to *spy* or to *watched*, and that this influences the interpretation of the sentence. The principle of minimal attachment means that the parser should attach the prepositional phrase to *spy*, because that is the current phrase. Of course, if there are pragmatic or semantic reasons why this would create a nonsensical sentence, as in the following sentence, then in the second stage of the model, the structure is revised to attach high to the verb phrase.

17. The spy watched the bomb with the binoculars.

Frazier (1987) argued that late closure is helpful from a processing perspective because it has a lower processing cost. Keeping a phrase "open" to new words until it is no longer grammatical to do so involves a lower processing load than closing and then reopening the phrase.

The problem with this particular principle is that it does not appear to be true for all languages. In English, there is a documented preference for "low attachment," but in other languages, like Spanish, the preference appears to be the opposite—attaching high. Consider the following two sentences from Cuetos and Mitchell (1988).

18. The journalist interviewed the daughter of the colonel who had had the accident.
19. El periodista entrevisto a la hija del coronel que tuvo el accidente.

Cuetos and Mitchell found while English speakers generally preferred an interpretation of (18) in which it was the colonel (and not the daughter) who was involved in the accident, Spanish speakers preferred the opposite in a sentence like (19). Further, eye-tracking data showed that the preference for high attachment in Spanish was present early in the sentence and so was not the result of a process in which late closure first applied and then was overridden.

This is not limited to Spanish (or English)—some languages appear to take one approach while other languages take the other. What does this mean for the principle of late closure? Because it doesn't apply in all languages, it is difficult to maintain that it is a universal or general processing strategy. Nonetheless, it could be argued that it is still the most efficient strategy for the parser, and the default. In this case, some languages, for whatever reason, are exceptions to this processing default. Another possibility, and one advocated by proponents of constraint-based models, is that what looks like a late closure preference is simply a reflection of the statistical frequency of certain structures in the language. So, in English, the final interpretation of a sentence involving a low attachment is more frequent than an interpretation with high attachment, and so comprehenders come to expect and anticipate low attachment. Conversely, speakers of Spanish come to anticipate high attachment because that is the structure more frequently encountered in Spanish. According to this argument, there really isn't a late closure principle at all, but instead the appearance of one under some circumstances based on a larger set of (semantic, syntactic, pragmatic) information that is being used to determine the sentence structure. On that note, let's turn to the constraint-based approach.

Let's Show Some Constraint

Constraint-based models take a very different approach to how sentences are initially parsed and how mistakes are sometimes made. Instead of applying general principles regardless of context (which sometimes fail to give the correct interpretation), many types of information are used to incorporate an incoming word into a sentence structure. In the case of ambiguous cases, all possibilities are considered and ranked according to their likelihood, again based not only on syntactic information, but things like the frequency of occurrence for a particular structure and the semantic fit between the rest of the sentence and the current word. Mistakes can still be made, particularly when evidence from these sources is pointing in one direction but the ultimate correct structure is something else. To get a sense of this, let's go back and compare a couple of our earlier examples, repeated in the following:

20. The horse raced past the barn fell.
21. The glass submerged in the water cracked.

Why is it that (20) is so difficult while (21) is not? According to constraint-based models, this is because in the case of (20) there are a number of sources of information pointing toward a high likelihood that the verb "raced" is the main verb of the sentence and not the beginning of a reduce relative clause. First, there is the frequency with which the word "raced" appears as either a past tense verb or a past participle. In this sense, the word "raced" is ambiguous, and in fact some researchers have argued that syntactic ambiguities occur because of lexical ambiguities and that there is an important link between lexical and syntactic ambiguity resolution (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994). *Raced* occurs more frequently as a past tense form of race and so this may bias the processing system to rank the main verb interpretation higher than the past participle. Also, there is the semantic fit between *horse* and *raced*. As an animate thing, it is possible for a horse to do things (rather than having them only done to it), and further, horses are well known to race. Again, this causes the main verb interpretation to be ranked higher than the reduced relative. Compare this to (21): *submerged* is also ambiguous between a past tense and past participle form. However, in this case the fit between *glass* and *submerged*, unlike *horse* and *raced*, means that it is unlikely that the glass submerged something because the glass is an inanimate object and so in a literal context it is unlikely to be the thing doing the submerging. This type of factor can cause the reduced relative interpretation of the sentence to be ranked higher than the main verb interpretation, or at least have the reduced relative alternative be more readily available when the disambiguating *cracked* is encountered.

Importantly, verb-specific information is used in determining the likelihood of a particular syntactic structure. Another type of temporary ambiguity arises in English between whether a noun following a verb is a direct object or the beginning of a new sentence embedded within the main sentence. For example:

22. The student learned geometry over the summer.
23. The student learned geometry was hard over the summer.

In (22), there is a student who is learning a thing—geometry. In (23) there is a student who is learning a relation—geometry is hard. In the first case, the noun *geometry* is a direct object of the verb *learned*, while in the second case the noun is beginning of a whole sentence that could stand on its own as a complete thought: "Geometry is hard." According to the garden path model, the minimal attachment principle means that the simpler structure will be chosen initially—*geometry* is a direct object of the verb rather than the start of a whole new sentence. This will be true regardless of the verb. However, according to the constraint-based model, if a particular verb, such as *confess* or *realize*, does not frequently occur with a direct object, then this should influence the syntactic structure that is initially highly ranked (or chosen). In particular, if a verb does not usually occur with a direct object, then if a noun immediately follows the verb it may well be interpreted as the beginning of an embedded sentence. In short, according to the constraint-based model, the lexical information specific to particular verbs is used from the very beginning during sentence processing, while in the garden path model verb-specific information is only used in the second stage of processing.

Before continuing, an important reminder is in order. The autonomy seen in garden path models is limited to the first stage of processing, and so while it can be illustrative to consider sentences like (20) and (21), it is important to remember that we are most interested in the first moments of processing, which are below conscious awareness. In many cases sentences pose "difficulty" only to the extent that they cause comprehenders to consistently linger very briefly (perhaps just an extra 100 milliseconds) on a word compared to a control word in a similar sentence. The comprehender him- or herself might not notice that one sentence is more difficult than another. It is important to emphasize this, because it makes the job of distinguishing the two models more difficult. We are looking for differences that are below the threshold of conscious attention and that occur in the very first moments of encountering each word in a sentence. With this idea in mind, we can turn now to some of the experimental evidence regarding ambiguous sentences.

Garden-Path or Constraint-Based? The Evidence Weighs in

One of the reasons that this topic has lasted so long in psycholinguistics, that is, why researchers could not agree about the flow of information in how sentences are processed, is that the evidence has been unclear, with both sides claiming strong supporting evidence. This highlights both the difficulty of designing experiments to test for differences in the very earliest stages of processing sentences, and the importance of testing what one thinks one is testing. Part of the problem has been that because the models are really distinguished by how they initially deal with a word, it has been difficult to conclusively show that context has an effect on this earliest stage. Context clearly does have an effect, but in order to distinguish the models, it must have an effect on *initial* processing. One way that this has been tested is by looking at what happens when the reader encounters the word that disambiguates the sentence. The basic idea is that if this word disambiguates in favor of the initial interpretation, then reading times for this word should be as fast as an equivalent unambiguous version of the sentence. On the other hand, if the word disambiguates against the initial interpretation, then reading times should be longer because the reader will need to reassess or reconstruct the sentence structure to accommodate the new word. Rayner and Frazier (1987) found this type of effect in sentences like (24) and (25).

24. The criminal confessed that his sins harmed many people.
25. The criminal confessed his sins harmed many people.

In this case, (24) is the unambiguous control condition because it contains the word *that* which marks the beginning of the embedded sentence *his sins harmed many people*. The sentence in (25) should have the same structure, but because it does not contain the word *that*, it is briefly ambiguous: *his sins* could either be the beginning of an embedded sentence (which it is) or the direct object of *confessed*. The principle of minimal attachment predicts that readers should initially build a sentence structure in which *his sins* is a direct object. Rayner and Frazier (1987) found evidence that this was indeed the case: readers slowed down at *harmed* in (25) compared to (24). So, it is clear that readers can be led down the garden path to the wrong initial interpretation, the question now is—is this caused because autonomous processing principles are applied regardless of context or is it because context is sometimes misleading?

Some Evidence in Favor of Garden Path, Autonomous Models

Some studies have shown that prior semantic context does not appear to influence or prevent initial syntactic misanalysis (e.g., Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; van Gompel & Pickering, 1998). Ferreira and Clifton (1986) looked at sentences like the following:

26. The defendant examined by the lawyer turned out to be unreliable.
27. The evidence examined by the lawyer turned out to be unreliable.

Their materials were designed to test whether the animacy of the subject noun and its fit to the verb influenced whether people initially misinterpret the verb as being a main verb rather than a reduced relative clause. The garden path model predicts that people will still initially attach the verb *examined* as a main verb, even when it is a poor fit in terms of the semantics of the sentence. This means that readers should show processing difficulty beginning at the word *by*, when it becomes clear that the verb is in fact the beginning of a reduced relative clause. The constraint-based model predicts that only (26) should show such a garden path effect. Ferreira and Clifton (1986) found evidence for a garden path effect for both sentences, suggesting that readers in both cases initially interpreted the verb as the main verb of the sentence. Thus, their results supported the garden path model.

Further, Ferreira and Henderson (1990) looked at temporarily ambiguous sentences and their unambiguous counterparts like the following:

28. Bill hoped Jill arrived safely today.
29. Bill wrote Jill arrived safely today.
30. Bill hoped that Jill arrived safely today.
31. Bill wrote that Jill arrived safely today.

In (28) and (29), there is a temporary ambiguity due to the fact that the word *that* has been omitted—it is not immediately clear at the word *Jill* whether this word is the direct object of the verb (*hoped/wrote*) or the beginning of an embedded sentence. This is not the case in (30) and (31), in which the word *that* marks the beginning of the embedded sentence *Jill arrived safely.* However, notice that in (28) it would be exceedingly strange if not ungrammatical for the sentence to end immediately after *Jill*; one cannot really say *Bill hoped Jill.* According to the garden path model, the sentences are subject to minimal attachment: in (28) and (29) the noun *Jill* should initially be attached to the sentence structure as the direct object of the verb *hoped/wrote.*

However, if the verb-specific information about *hope* is relevant during initial processing, as it is according to the constraint-based model, then it should prevent readers from initially being led down the garden path: *Jill* should be attached as the beginning of a embedded sentence because it cannot reasonably be the direct object of *hoped*. However, Ferreira and Henderson found a garden-path effect for both (28) and (29) compared to their unambiguous counterparts. They used this finding to argue that the sentence parser does not initially take individual verb usage/bias into account and that such verb information cannot override minimal attachment.

Some Evidence in Favor of Constraint-Based, Interactionist Models

First, let's turn to some work done by John Trueswell and his colleagues that follows up the findings from Ferreira and Clifton (1986). Trueswell, Tanenhaus, and Garnsey (1994) contended that the materials that Ferreira and Clifton used did not effectively manipulate semantic fit; they argued that about half of the materials with inanimate nouns could still continue plausibly with a main verb instead of a reduced relative clause. For example, while inanimate, the sentence beginning "The car towed..." could still plausibly end with *towed* as a main verb instead of a reduced relative verb. Further, Trueswell et al. had concerns about the presentation of the sentences in the study—in Ferreira and Clifton (1986), it was necessary that the sentence be split across two lines. Ferreira and Clifton decided to split the sentence at the same point in the disambiguating region, but this meant that for the ambiguous reduced relative clauses, the presentation looked "unnatural," according to Trueswell et al. Further, they were skeptical about whether the unreduced relative clauses that Ferreira and Clifton used were really appropriate as a baseline control condition (recall this same type of concern from the lexical ambiguity research we discussed earlier). So, Trueswell and his colleagues conducted a study to attempt to replicate Ferreira and Clifton's while addressing these concerns. The sentences appeared all on one line, the inanimate versions could not continue with a main clause interpretation of the verb, and in addition to the control condition used by Ferreira and Clifton, Trueswell et al. added a second control condition in which the form of the verb made the sentence unambiguous (e.g., The poster drawn….). In contrast to Ferreira and Clifton, they found that the semantic fit of the subject with the verb had a reliable effect: garden path effects were reduced with inanimate nouns. Further, the degree of reduction was relative to the strength of the semantic fit (or lack thereof), according to a separate study that asked participants to rate how typical it is for things like evidence to examine something. These results cast doubts on the reliability of Ferreira and Clifton's results; however, in further work Clifton and colleagues challenged Trueswell et al.'s evidence and conclusions (Clifton et al., 2003). They, in turn, had concerns about the presentation of the sentences in Trueswell et al.'s study. In particular, they argued that because participants could see the disambiguating word in advance of fixating on it (called a preview effect), reanalysis processes could be triggered before encountering the disambiguating word and thus artificially diminish the effect. That is, inanimate subjects do not prevent garden path effects, instead participants in the Trueswell et al. study began the reanalysis process sooner than they thought. Clifton et al. (2003) provided some experimental results to support this argument.

Other work has examined whether verb-specific information can have an early impact on sentence processing. While some work (e.g., Ferreira & Henderson, 1990) has shown that verb information does not make a difference, other research has shown the opposite. Garnsey, Pearlmutter, Myers, and Lotocky (1997) examined temporary ambiguities involving verbs that could either be followed by a direct object or an embedded sentence. However, first, they conducted a preliminary study to determine the way that the verbs are normally used: they divided verbs into those that people usually followed with a direct object (DO-bias verbs), those that were usually followed by an embedded sentence (also known as a sentence complement, SC-bias verbs), and those that occurred with both types of continuations roughly the same number of times (EQ-bias verbs). They then created materials in which the noun following the verb was either a plausible or implausible object of the verb. Examples of the beginnings of sentences from the study are given below, with the critical verb underlined and then followed by the plausible and implausible object nouns.

32. DO-bias: The talented photographer *accepted* the money/fire…
33. SC-bias: The ticket agent *admitted* the mistake/airplane…
34. EQ-bias: The proud mother *announced* the wedding/flowers…

Garnsey and her colleagues found that the bias of the verb had an impact on their results: While there was a plausibility effect for DO-biased verbs in which words like *fire* were read more slowly than *money*, no such effect was found for SC-bias verbs. This is consistent with using verb-specific information to aid in structure building but is inconsistent with autonomous models in the verb bias would not taken into account. These results thus provide evidence in favor of an interactionist, constraint-based model of sentence processing.

This brief overview of these experiments should give us a sense of the difficulty in pinning down the flow of information during sentence processing. The importance of the timing of any influence of non-syntactic context is vital to distinguishing these models and understanding when semantic and other information is available during sentence processing. Reading studies alone have not provided an entire picture, but fortunately there is also evidence from another source: a technique called the visual world paradigm that we discussed in chapter 3. Recall that in this technique, participants listen to language stimuli while viewing a visual context, such as an array of objects or pictures of objects. The eye movements of the participants are recorded and time-locked to the language that they hear. This technique is relevant here because it can show whether and when participants use visual information during sentence processing. A number of studies have shown that visual information is used immediately and even predictively during sentence processing.

First, let's revisit the results from Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy (1995) that we discussed in chapter 3. In chapter 3, the importance of these results was to establish how the visual world paradigm works, but these results are also important with respect to ambiguity resolution. Recall that participants heard sentences like (35) and (36) while looking at a monitor that displayed a set of four pictures. In one scene, the pictures were an apple sitting on a towel, an empty towel, a pencil, and an empty box. In the other scene, the pencil was replaced by an apple sitting on a napkin—meaning that there were two apples shown.

35. Put the apple on the towel in the box.
36. Put the apple that's on the towel in the box.

In (35), it is ambiguous whether "on the towel" attaches low to modify apple, or whether it attaches high, to state where to put the apple. The visual scenes were designed to bias the attachment in a particular direction: the scene with only one apple should cause a bias toward interpreting *on the towel* as specifying a location to move the apple to. This is because there is only one apple and so no need to further describe it. On the other hand, the two-apple scene created a bias toward interpreting *on the towel* as a description of the apple (distinguishing it from the apple on the napkin). Tanenhaus et al. (1995) found that these two scenes, despite being different only by an apple, caused very different eye movement patterns: After hearing *apple* in the one-apple scene, participants looked first to the apple and then to the empty towel, reflecting an interpretation of towel as the destination. After hearing *apple* in the two-apple scene, participants looked at both apples but did not look to the empty towel. This is consistent with interpreting *on the towel* as a description of the apple. Importantly, these results show that visual context can alter attachment preferences, which an autonomous model of sentence processing would not predict.

Even stronger evidence that visual information influences sentence processing comes from studies of non-ambiguous sentences. In some cases, it appears that visual information not only helps guide parsing but allows comprehenders to predict upcoming structure. One study by Kamide, Altmann, and Haywood (2003) used a visual world paradigm design to examine the interaction of visual information and verb-specific information in sentences with ditransitive verbs. Ditransitive verbs, like *put*, *slide*, and *spread*, often or necessarily occur with two objects or an direct and indirect object. For example, in (37), the verb *put* requires both *the book* and *on the table* in order to be grammatical.

37. Bob put the book on the table.
    *Bob put the book.
    *Bob put on the table.

Kamide et al. combined sentences like *The woman will slide the butter to the man* with a visual scene of a man, woman, butter, and a loaf of bread. When participants heard the sentence with *slide* as the verb, they began to look at the man shortly after hearing *butter*. This shows that participants anticipated the next noun of the sentence based on the visual scene plus information from the verb *slide*. They compared this with sentences like *The woman will spread the butter on the bread*. The only difference before *butter* is the verb *spread*, and yet in this case participants reliably looked to the loaf of bread more than the man. These results provide evidence that people can rapidly combine verb-specific information with visual information to anticipate upcoming sentence structure.

WRAPPING UP

It seems clear, not to mention intuitively satisfying, that things like plausibility and frequency play a role in how we interpret sentences when they are ambiguous. The key question is how quickly they are able to have an influence. Two major approaches have proposed different solutions to how to make language processing both quick and accurate, while also allowing errors to occur. At the moment, the balance of evidence appears to favor an interactionist approach for sentence processing, while for words it is less clear whether context can override frequency of meaning. However, for words the question is not whether context has an effect, but how strong an effect. With respect to sentences, there is a new approach that we will explore in chapter 8 that questions the extent to which we even build complete syntactic structures as we comprehend sentences.

REFERENCES

Clifton, C., Traxler, M., Mohamed, M., Williams, R., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory & Language*, *49*, 317–334.

Cuetos, F., & Mitchell, D. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, *30*(1), 73–105.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory & Language*, *25*(3), 348–368.

Ferreira, F., & Henderson, J. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*(4), 555–568.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586). Hove, UK: Lawrence Erlbaum Associates Ltd.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.

Garnsey, S., Pearlmutter, N., Myers, E., & Lotocky, M. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory & Language, 37*, 58–93.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory & Language, 49*, 133–156.

MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). Syntactic ambiguity resolution as lexical ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 123–153). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Martin, C., Vu, H., Kellas, G., & Metcalf, K. (1999). Strength of discourse context as a determinant of the subordinate bias effect. *Quarterly Journal of Experimental Psychology, 52A*(4).

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition, 9*(3), 225–236.

Rayner, K., Binder, K. S., & Duffy, S. A. (1999). Contextual strength and subordinate bias effect: Comment on Martin, Vu, Kellas, and Metcalf. *Quarterly Journal of Experimental Psychology, 52A*(4), 841–852.

Rayner, K., & Frazier, L. (1987). Parsing temporarily ambiguous complements. *Quarterly Journal of Experimental Psychology, 39A*, 657–673.

Sereno, S. C., O'Donnell, P. J., & Rayner, K. (2006). Eye movements and lexical ambiguity resolution: Investigating the subordinate-bias effect. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 335–350.

Tanenhaus, M. K., Spivery-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.

Trueswell, J., Tanenhaus, M. K., & Garnsey, S. (1994). Semantic influences of parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory & Language, 33*, 285–318.

van Gompel, R., & Pickering, M. J. (1998). Lexical guidance in sentence processing: A note on Adams, Clifton and Mitchell. *Psychonomic Bulletin & Review, 8*(4), 851–857.

5

(Multiple) Language Representation and the Brain

In addition to an interest in understanding how linguistic and (nonlinguistic) information interacts during language processing, researchers are also interested in how language is represented, not only in terms of where in the brain language functions are located but also in terms of the mental constructs of language. In this chapter, we'll first talk about the representation of language in the brain—including what parts of the brain are known to be involved in language. Then, we'll spend the rest of the chapter talking about how multiple languages are represented and interact in bilingual speakers.

LANGUAGE AND THE BRAIN

Let's start with a quick and light overview of brain anatomy. First, the main part of the brain that is relevant for our purposes is the cerebrum or neocortex, a thin layer of tissue that forms the outermost surface of the brain. In human brains, this part has a very characteristic bumpy appearance, with lots of peaks and troughs. This part of the brain is responsible for all kinds of fundamental aspects of cognition, motor movement, and sensation. For example, the cerebrum contains primary motor cortex, which is responsible for intentional body movements, and visual cortex, which is responsible for processing information from the eyes and allowing us to recognize the identity, location, and movement of objects. The cerebrum is divided into four lobes and each of these lobes in turn has regions that are known to be important for particular functions. The most important lobes for language are the temporal lobe (near the ear) and the frontal lobe (at the front of the brain). Language function appears to be heavily dependent on certain areas within these lobes, but there are also areas deep within the brain (called subcortical areas) that are involved as well.

Further, the brain is divided into two halves, or *hemispheres*. These two hemispheres are not exactly mirror images of each other in terms of the functions that they support—instead many higher-level cognitive functions are lateralized, meaning that parts of the right hemisphere and left hemisphere specialize and take on functions that the other hemisphere does not. There are a few exceptions, but for the most part, the brain is organized such that the left hemisphere controls volitional movement and processes incoming sensory information from the right side of the body and the right hemisphere is responsible for these same functions but on the left side of the body. In terms of language, in right-handed people it is the *left* hemisphere that supports the majority of language function. In left-handed people it varies somewhat.

Most left-handers also have language dominance in the left hemisphere, although for some, language may actually have diffuse representation across the hemispheres or even have a right hemisphere dominance. However, even in right-handed people, the right hemisphere also plays a role in language ability, although exactly what this role is and how it works are still something of a mystery.

There are two areas in particular that appear to be especially important for language: an area toward the front of the brain in the frontal lobe that includes Broca's area and an area more or less beneath and behind the ear toward the back of the temporal lobe called Wernicke's area. See Figure 5.1 for a schematic view of their positions in the cortex.

Classically, Broca's area has been associated with speech production and Wernicke's area with auditory comprehension of speech sounds. This is in part due to the constellation of impairments that appear in either the production or comprehension of language (or indeed both) when patients suffer brain damage in these particular areas. A major source of data on the organization of language in the brain comes from studying

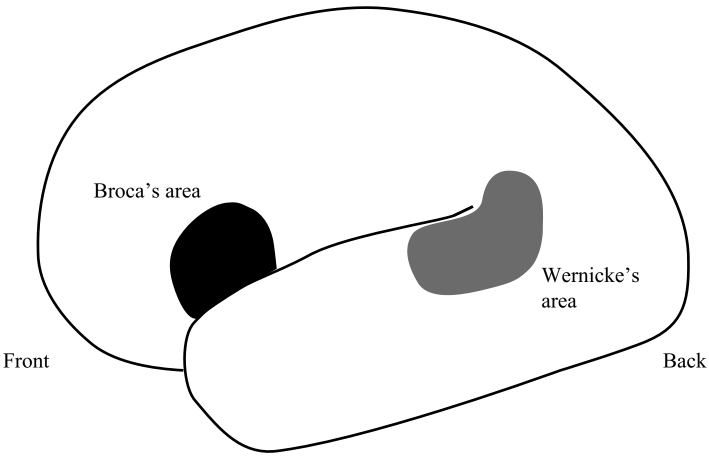

FIGURE 5.1 Key language areas in the brain.

patients with aphasia, a language disorder caused by damage to the brain.

There are many kinds of aphasias, including Broca's aphasia—which is caused by damage to Broca's area as well as to some adjacent brain areas, particularly toward the front of the brain—and Wernicke's aphasia, which is caused by damage to Wernicke's area. Broca's aphasia is characterized by difficulty with language production—with effortful, slow speech, and the striking absence of function words like prepositions, determiners, conjunctions, and grammatical inflections. So, the majority of speech in this type of aphasia is composed of nouns and verbs, produced without a fluent connection between them. Despite their often severe problems with language expression, patients with Broca's aphasia do appear to have largely intact comprehension skills and the content that they do produce is on topic and the intended meaning is often very clear. This kind of aphasia contrasts sharply with the pattern of language loss seen when there is damage to the back of the temporal lobe (i.e., Wernicke's aphasia).

Wernicke's aphasia is, in many ways, the opposite of Broca's—patients with this aphasia speak fluently and have no trouble with function words. However, the content of the speech is often not meaningful and may even contain word-like strings of sounds that are not actually words (aphasia researchers call these neologisms). Or, they may produce novel ways to refer to things, such as calling an egg "hen-fruit." In terms of comprehension, they often show clear signs of auditory comprehension difficulties and they commonly have great difficulties repeating spoken words.

What does this mean for language representation in the brain? While Broca's area was originally seen as important primarily for producing language, researchers have more recently proposed that Broca's area is crucially involved in our ability for structuring language—thus crucially involved in syntax. In a landmark paper on this topic, Caramazza and Zurif (1976) found that patients with damage to Broca's area not only had problems with producing sentences but also had some comprehension problems as well: they had problems with getting the meaning right for sentences when the use of syntactic information was required. For example, in a sentence like "The book was read by the girl," one does not really need syntax to understand the event that is being described—one only needs to have real-world knowledge that girls read books but books cannot read girls. However, if instead, you had a sentence like "The boy was kissed by the girl," real-world knowledge alone doesn't help—boys can both kiss and be kissed, and the same is true of girls. So, who kissed whom? The syntactic structure and corresponding rules of English tell us that it was the girl who did the kissing because this sentence is a passive—indicated by the word "was" and the "ed" on the verb, as well as the prepositional phrase introduced with "by," which gives the agent of the kissing action. However, in active sentences like "The girl kissed the boy" the subject of the sentence is the one who is the agent. In fact, as we'll see in chapter 8, there is a strong preference to interpret the subject of a sentence as the agent of the verb. So, what Caramazza and Zurif found was that patients with damage to Broca's area not only had problems with language production but also had difficulty figuring out that it was the girl, and not the boy, who did the kissing in passive sentences where the roles of kisser and kissee could be reversed. However, other work (Linebarger, Schwartz, & Saffran, 1983) found that patients with Broca's aphasia were actually not so bad at recognizing ungrammatical sentences, suggesting that they do have some preserved syntactic ability.

One of the problems with looking at patient data is that there is an enormous amount of variation—no two patients diagnosed with Broca's aphasia are exactly alike. This has made some researchers call into question whether Broca's aphasia is really a single problem that is useful as a diagnosis (e.g., Caramazza, Capitani, Rey, & Berndt, 2000) and it has certainly added to confusion about how to diagnose and group patients. However, one large-scale study of Broca's patient data revealed that there are indeed some key impairments that Broca's patients share (Drai & Grodzinksy, 2006). In particular, one of the key difficulties with respect to syntactic processing appears to be dealing with noncanonical sentences—sentences in which words are out of their usual order, such as passives (like those from Caramazza & Zurif, 1976), but also sentences with object-relative clauses, such as "The man knew the boy who the girl kissed" in which *the boy* is the object of the verb *kissed*. This analysis confirms what many individual studies have also found across different types of tasks.

In one recent study, Dickey and Thompson (2009) examined eye movements of patients with Broca's aphasia while they listened to short passages that used passive voice or object-relative clauses in a critical, final sentence. An example of their materials is given next:

1. One day a bride and groom were walking in the mall.
   The bride was feeling playful, so the bride tickled the groom.
   A clerk was amused.
   Point to who the bride was tickling in the mall.

While listening to these sentences, participants would look at a computer screen that displayed four pictures related to the passage, for example, a groom, a bride, a mall, and a clerk. At the end of the passage the participants would point to a picture. Dickey and Thompson compared eye gaze patterns from the patients with unimpaired participants and found that while the eye movements between the groups were largely similar at the beginning of sentences, they diverged by the end and that the patients took longer to look at the correct picture (e.g., groom) and appeared to be more distracted by the grammatically impossible subject of the critical sentence (e.g., bride). These results and many others confirm that in Broca's aphasia one key deficit is in understanding grammatical structures that involve noncanonical word ordering, particularly displacing grammatical or logical objects to earlier positions.

Wernicke's aphasia has a very different pattern of deficits from Broca's aphasia. Patients with Broca's aphasia often struggle to produce language, and show clear grammatical deficits, whereas patients with Wernicke's aphasia can produce fluent speech with plenty of grammatical morphemes. However, the striking feature of speech from patients with Wernicke's aphasia is lack of semantically meaningful context. It is not uncommon in Wernicke's aphasia to see nonsense words used as if they had meaning (remember neologisms). In terms of comprehension, Wernicke's patients generally show clearer deficits than in Broca's aphasia, with many patients showing signs of difficulty in understanding even very straightforward sentences with all words in their usual places.

Broca's and Wernicke's aphasia are not the only types of aphasia. In fact there are several others, including resulting deficits from damage to the connection between Broca's and Wernicke's areas (i.e., conduction aphasias) as well as damage to connections to association areas from these areas (i.e., transcortical aphasias). Disordered language also arises not only from acute brain injuries but also from degenerative brain disorders like Alzheimer's disease and frontotemporal dementia. One specific form of frontotemporal dementia is semantic dementia (SD). As the name suggests, the key deficit in this disorder is a progressive decrease in functioning semantic knowledge (Snowden, Goulding, & Neary, 1989). This manifests initially as having difficulties in dealing with the meanings of words—both in production and comprehension contexts. For example, patients will often produce what are called semantic "paraphasias" when they are asked to name pictures, in which they provide a (somewhat) related word instead of the intended or target word (such as calling an apple a "ball"). Indeed, naming difficulties are a key feature of SD and have been shown to reflect a deteriorating system of semantic knowledge—in particular, there is evidence that as SD progresses, patients lose knowledge that allows them to specify between similar concepts—so that a cat and an elephant may both be simply "animal" (Reilly & Peelle, 2008).

At the same time, patients with SD may produce fairly fluent discourse in terms of speech rate and may not even show any significant impairment to a casual observer under typical conversational circumstances. However, a more careful look at their discourse shows that, consistent with their increasing deficit in semantic knowledge, SD patients produce fewer low-frequency words, and that this pattern gets stronger as the disease progresses. The overall level of meaningful content also declines with severity, while there is greater use of more generic terms (e.g., thing, animal) and highly frequent functional words (like "the" and "this"). On the other hand, phonology remains largely intact even until relatively advanced stages of the disease.

The underlying cause of SD provides some insight into how it is that we know what things are and the interaction of this knowledge with language function. SD is a neurodegenerative disease, meaning that it involves the deterioration of neurons. In SD, this means a progressive deterioration in the temporal lobe that typically (though not exclusively) begins in the front (anterior) of the lobe and then spreads back (posterior), generally with a larger impact on the left hemisphere than the right (Galton et al., 2001). The progressive nature of the anatomical changes in the brain is reflected in the progressive disruption of semantic ability. However, unfortunately we cannot draw strong conclusions between atrophy in one area and subsequent semantic disruption. Nonetheless, studies of the progression of SD and the nature of the neurodegeneration underlying it allow us to consider neurological evidence regarding the nature of semantic representation in the brain.

Reilly and Peelle (2008) discuss the two basic approaches to semantic representations as they apply to SD: on one hand there are number of theories that are distributed or embodied in nature. In these theories, semantic concepts are decomposed into different features that are stored in modality-specific areas (such that visual information about the concept would be stored in areas related to visual processing, auditory information would be stored in areas related to auditory processing, and so on). Activation of a "single" concept would actually entail coordinated activation of the decomposed features, spread out over different areas of the cortex. There is some empirical evidence for this: a number of different studies have found that activating the concept of something also creates activation in modality-specific areas (e.g., pictures of tools activate areas of the cortex related to the motor coordination in grasping) (Pulvermüller, 2005). On the other hand, there is also an "amodal" approach in which there is a central area where semantic memory is stored rather than a distributed system alone. Reilly and Peelle proposed a hybrid model in which information from different modalities converge on single representations, but that this central representation lacks information about sensory details—these are stored in a distributed fashion. The central representation is a kind of index in this regard—it holds very little information of its own, but serves to link to richer, distributed sources of information. In Reilly and Peelle's approach, impairment in SD comes in two parts: impairment to abstract semantic representations and impairment to visual semantic features. (Interestingly, SD also impairs the ability to recognize faces and objects.)

The fact that deficits in semantic knowledge can manifest with other deficits brings us nicely to another topic to consider when thinking about language and the brain: what other cognitive functions are important for language, and so might affect language ability when disrupted? Working memory is one such function, and it has been widely studied both in its own right as well as in the context of language processing. The other function is executive function, which is less well studied in the context of language, but has been increasingly identified as important to higher level language functions.

In general terms, working memory is a short of mental scratch pad—a space in which information is temporarily stored and operations on that information are carried out. One influential model of working memory (e.g., Baddeley, 1992) is principally composed of three parts: a visuospatial sketch pad (temporarily storing visual information), a central executive (directing attention), and a phonological loop (in which phonological information is held). In this model, the central executive and the phonological loop play a critical role in language, with the central executive involved in semantic integration and the phonological loop involved in phonological processing. This is not the only view of working memory; indeed, there are many theories about how working memory works and how it relates language. One common task that is used to access working memory capacity as it relates to language is the reading span task (Daneman & Carpenter, 1980). In this task, people are asked to read sentences aloud and then recall the last word of each sentence. If this were done one sentence at a time, it would not be very challenging. The key aspect of this task is that participants must keep several words in memory while still reading aloud. So, at the beginning of the task, a participant will read two sentences aloud and then report the last words of those two sentences, in order. They will do this for several pairs of sentences, and then be asked to do the same thing but with sets of three sentences, then with sets of four, five, and finally six sentences.

Using this task, and other tests of working memory capacity, researchers have found that a person's working memory capacity has an influence on sentence processing, and in particular on syntactic processes. King and Just (1991) found, for example, that participants with lower reading spans showed greater difficulty with object-relative clauses than participants with higher reading spans. For example:

2. The reporter who the senator harshly attacked admitted the error.

Working memory is known to be especially relevant for sentences like these that involve displaced constituents. Wh-questions are another example—in both wh-questions and relative clauses, comprehenders must associate some word (called a filler) with a position elsewhere in the sentence (called a gap). This is shown in the following wh-questions, with the filler word in bold and the gap represented by an underline:

3. **Who** did the woman drive ____ to the airport.
4. **Who** ______ drove the guest to the airport?

In (3), the wh-word *who* must be linked in some way to the gap following the verb *drive* because it is the person who was driven that is questioned. By the same token, in (4), the word *who* must be linked to the gap immediately preceding the verb. There is an interesting asymmetry for both relative clauses and wh-questions in which it appears to be more difficult to make this link when the filler is an object (as in (3)) compared with when it is a subject (as in (4)), although there is some recent work suggesting that this is not always the case (Roland, O'Meara, Yun, & Mauner, 2008).

In this section, we have covered a number of topics related to how language is represented in the brain, including a brief overview of the areas important for language, aphasias, and other language disorders, and some nonlinguistic functions that nonetheless contribute to language processing, particularly working memory. In the rest of this chapter, we will consider the case of multiple languages.

MULTIPLE LANGUAGES

Since we are still not in complete agreement about how a *single* language might be represented in the brain, it might seem like jumping the gun a bit to start worrying about how *two* languages work, let alone three or more. However, bilingualism is the norm world wide, with at least 50% of the world's population speaking two or more languages (some estimates go as high as 70%) and so it makes sense to see how multiple languages are represented. Not only are we then studying the usual state of affairs, but information from bilingual representations can inform monolingual (single-language) representations as well.

Perhaps the biggest question in multiple language representations is how the languages interact and overlap in terms of representation and to what extent they are held distinct. That is the question that we'll look at in the rest of this chapter.

First, however, we need to address a few terminological and methodological issues. First, and most pressing, what is a bilingual? If suddenly teleported into the middle of France, I could probably manage to get by with my increasingly rusty knowledge of French. I can read a French newspaper (with a certain amount of guesswork or a dictionary), and I can understand parents speaking in French to their young children pretty easily. Does that make me bilingual?

According to some definitions, it would. I am, without question, a second language learner of French because I didn't begin to learn the language until I was 14. But, what if I had started when I was 7? Clearly, we need to acknowledge that there are different types of bilingualism. A *native* bilingual is someone who began learning more than one language from a very early age. A related idea to this is being a simultaneous bilingual, which is someone who learns a second (or third or fourth) language without first having mastered a first. A simultaneous bilingual is also a native bilingual, but the precise age for "native" status is a little more tricky, and so simultaneous is perhaps the better term because it is more precise. A sequential bilingual is someone who learns a second (or third, or fourth) language after achieving relative mastery of a first language. So, even someone who begins to learn a language at the age of 7 is a sequential bilingual because they will by that time already have a first language that they have mastered.

Another issue for studying bilingualism is acknowledging that it is extremely rare that a bilingual speaker would have equal strength in all areas of all their languages at the same time. For example, even if a bilingual speaker uses both languages equally on a daily basis (which is, itself, somewhat unusual), they are likely to have certain topics or uses of language that are language-specific. So, relative vocabulary strengths can differ between languages quite easily: if you only use one language to talk about your work in, then this language has an advantage over your other languages in this topic. Many bilingual speakers do not live in a balanced language environment: one language will be used relatively more than the other, and in this case, the less used language may be at a disadvantage. Further, while someone may be able to speak more than one language, there may be political or social reasons why they use one as little as possible, or have a poor attitude toward it. This, too, can influence proficiency in the language. Taken altogether, these observations highlight that there is a lot of individual variation among bilingual speakers, even when they appear from certain measures to be identical. For example, you could have a group of 30 French–English bilinguals who all began learning English between the ages of 7 and 10 and are currently between the ages of 18 and 25. Despite this similarity, they may still have very different skills in English (and possibly even French), depending on whether they are currently living in an English-dominant, French-dominant, or balanced bilingual environment. So, proficiency is an important thing to assess in studies of bilingualism, although no single test will be able to assess all aspects of language proficiency.

THE BILINGUAL BRAIN

A great deal of work has looked at bilingual language processing (for an overview, see Nicol, 2001). In terms of where multiple languages are represented in the brain, there are several studies using fMRI and PET that show that multiple languages activate largely overlapping regions (e.g., Chee et al., 1999; Illes et al., 1999; Perani et al., 1998). However, these studies generally report averages across participants rather than patterns within individuals. Evidence from bilingual aphasia patients clearly shows that the relationship between languages at the neurological level is more complicated than "largely overlapping" and highlights the fact that there is a lot of individual variation in language representation. For example, bilingual patients with aphasia may suffer language loss in only one language, and may even have language loss in their native language while retaining function in a second language. They may retain both languages, but have difficulty translating between them, or they may switch uncontrollably between them. They may be able to translate into a language that they can no longer speak spontaneously. And, of course, they may suffer language ability losses in both languages, to same or differing degrees. Just as the patterns of loss vary widely among patients, so do patterns of recovery. Patients may recover both languages at an equal rate and to an equal level, or only one language may recover. Further, both languages may ultimately recover to similar degrees, but may not recover at similar rates. Finally, and most intriguing, some patients have shown an antagonistic recovery pattern, in which gains in recovery for one language appear to come at the loss of recovery in another.

One fMRI study that looked at healthy bilinguals compared them at an individual level to examine which brain areas showed increased activity for first versus second language (Dehaene et al., 1997). They played short passages to French–English bilinguals that were either in French or English as well as a control condition that was a short passage in Japanese that was played backward (so as to subtract away activation that was due only to hearing language-like sounds). First, they found that there was a great deal of individual variation between speakers in exactly which areas of the brain became more active during listening, even in the first language, French. However, for the French passages, all participants did show activation in certain regions in the temporal lobe and near Broca's area. That is, while the exact areas differed somewhat, activation for French was relatively tightly clustered for each participant, and across participants, similar areas were activated. However, for English materials, there was a much larger variation in which parts of the cortex became active. While some traditional language areas showed activation, the activated areas were also more broadly distributed and two of the (right-handed) participants showed activation in the right hemisphere only. So, although, this showed that there is, in fact, a certain amount of variability between speakers even in their first language, the areas activated for the first language were more consistent and focal compared with the more widely distributed and more variable second language areas.

Another study (Perani et al., 1998) showed that proficiency may be more important than age of acquisition in determining how multiple languages are represented in the brain. The researchers found no difference in patterns of activation (now looking again at averages rather than individual participants) between the L1 and L2 of participants when those participants were equally proficient but had begun acquiring the L2 early or late. However, speakers who had begun acquiring their L2 at the same time, but had achieved different levels of proficiency, did show differences in the patterns of activation.

These studies, and other similar ones, show us several important things. First, it appears there is individual variation in precisely which parts of the brain become active during the same language stimuli. Next, second languages are not necessarily represented in the same areas as first languages, but they appear to become increasingly overlapping with first languages as people become more proficient. While earlier ages of acquisition are associated with increased proficiency, it may be proficiency itself rather than age of acquisition per se that influences the representation of language in the brain.

PSYCHOLINGUISTIC STUDIES OF BILINGUALS

Another perspective on the degree of overlap and interaction between languages comes from psycholinguistic studies of language function, rather than neurological representation. We can divide work from this perspective into three main sections: work looking at shared/distinction representations of phonology, lexicon, and syntax.

Studies in this area tend to follow a standard pattern: take an established finding from monolingual speakers, or at least from a single-language testing environment, and add a second language to the task to see how this addition influences processing compared with the single-language version.

Phonological Representations

Work on sound in bilingual speakers has mostly focused on the role of phonological information during word retrieval, and thus is really part of the lexical representation of the word. However, at least one study has looked at the nature of the representation of phonemes themselves in bilingual speakers. This is of interest because different languages have different sound systems—they carve up the sound space in different ways, as we discussed in chapter 2. For example, aspiration, the puff of air that is released when a stop consonant ends, may be extremely important in defining speech sounds, as it is in Hindi, or not important at all, as in French. So, what happens when you have multiple sound systems? Are similar sounds between the two (or more) languages shared?

There is some experimental evidence to suggest that similar sounds may indeed share some degree of representation. Roelofs (2003) investigated this using a task in which speakers memorize pairs of semantically related words, such as fruit–melon, iron–metal, and grass–meadow. Then, speakers asked to say aloud the second word of the pair when they see the first word. So, if the speakers saw the word *fruit* then they should say *melon*. Participants in Roelofs study saw sets of three word pairs. In some of these word-pair sets, the first syllable was the same across all three of the second words in the pair (e.g., melon, metal, and meadow). In other cases, words began with different syllables. Roelofs tested Dutch–English bilinguals and found that the sets in which the words all shared an initial syllable were responded to faster than the sets without a shared syllable. Importantly, this was true even when the sets contained words from both English and Dutch. So, a word set like *fashion–style*, *tafel–stoel*, *pebble–stone* was still faster than a set like *level–floor*, *trommel–blik*, *fashion–style*, even though the words with the shared syllables came from different languages. These results suggest that bilinguals may have a shared system of representations of phonological segments.

Lexical Representations

From the perspective of word representations, one of the big questions for bilingual speakers is to what extent representations of words from one language become active when the other language is used. More specifically, we can ask whether words from one language influence the processing of words in the other. A lot of work in this area has focused on speech production as opposed to comprehension, and that is what we will focus on in this section.

A typical task in research in this area is to ask speakers to name something in one language while also viewing a distractor word in the other, nonresponse language. Thus, speakers need to suppress information from one language while speaking in the other. The question is: what makes this suppression easier or harder? A particularly famous example of this type of task is the Stroop task, which has been widely used in single-language studies. In this task, people are asked to name the color that a word is presented in, ignoring the meaning of the word itself. So, if you see the word "table" displayed in blue, you would respond "blue" (not "table"). This is a seemingly trivial task and people are generally very, very accurate when doing it. But, the time it takes them to name the color can be reliably influenced by the meaning of the word: we cannot suppress retrieving the meaning of the word, even when it is not relevant to the task. When the word meaning is the same as the color it is printed in ("blue" presented in a blue font), then people are faster compared with when the word meaning is unrelated (as in the earlier "table" example). However, speakers are slower to name the color when the word meaning is a conflicting color term ("red" presented in a blue font). The meaning of the word interferes with the naming of the color, and we get two effects: an identity effect, in which times are faster when the word meaning matches the color to be named, and a competitor effect, in which times are slower when the word meaning is a competing color term. It turns out that this type of effect isn't limited to color terms: in a picture–word interference task people are asked to name a depicted object (e.g., a drawing of a table) while simultaneously seeing a distractor word superimposed over the picture (e.g., the word table, chair, or car). The idea is that the participant must ignore the word and name the object that's depicted. Just as in the Stroop task, the meaning of the word matters: there is an identity effect when the distractor matches the name of the picture and there is a competitor effect when the word is related, but different.

So, what happens when a bilingual speaker who is native or highly proficient in both languages has to name the picture in one language when the distractor word is printed in another—does the nonresponse language interfere with or aid the response language? There are three effects that have been found. First, there is a cross-language competitor effect in which naming times are slower when the distractor word is a related word, just as in single-language studies. So, if a Spanish–English bilingual is asked to name a picture of a table in Spanish, but the word "chair" is printed over it, they are slower than when the word "car" is printed over it. Second, there is a cross-language identity effect

(Caramazza & Costa, 1999; Hermans, 2004) in which participants name the picture faster when the distractor word is also the name of the picture, but in the other language. So, naming times for table in Spanish are faster if the distractor word is "table." Finally, there is another effect, called the phonotranslation effect, in which naming times are slower when the distractor word is phonologically similar to the picture name. This is true when the distractor word is phonologically similar in the nonresponse language. So, for example, let's say now that the picture to be named is a dog and it should be named in Spanish. So, the phonotranslation effect occurs when a word like "doll" is superimposed over the picture. This seems somewhat strange at first, but the idea is that the word "doll" activates the similar sounding "dog," which then acts like a competitor with the word "perro" (Spanish for dog). A similar effect also happens when the distractor word is phonologically related to the word to be named, but in the other language. For example, in one study of Dutch–English bilinguals, Hermans, Bongaerts, De Bot, and Schreuder (1998) found that when the speakers were asked to name pictures in English, having a English word that was phonologically similar to the (completely unspoken) Dutch name for the picture caused interference.

However, these phonotranslation effects may be limited to picture–word interference tasks: a recent study (Costa, Albareda, & Santesteban, 2008) found either no difference or *faster* naming times in a Stroop task version testing this effect. Taken together, these results confirm that during lexical access the lexical representations of both languages become activated at the same time. However, the relative degree of interference between the representations is still unclear—with some studies showing interference and others not.

Syntax

Recall that the outstanding question here is to what extent the representations of the two (or more) languages are integrated into a shared system and to what extent they are held distinct. One very useful technique for exploring to what extent the grammars of two languages overlap is syntactic priming. Under certain circumstances, monolingual speakers are more likely to produce a sentence with the same syntactic structure that they have recently heard. So, for example, as shown in Bock (1986), if a speaker hears a sentence like "The ball was kicked by the boy" and is then asked to describe a picture that depicts a church with a tall steeple and a bolt of lightning such that the bolt of lightning is in contact with the steeple, speakers are more likely to produce a sentence like "The church was struck by lightning" than "Lightning struck the church." However, if the speaker heard "The boy kicked the ball," then the reverse is true—"Lightning struck the church" is more likely. This issue of syntactic (or structural) priming will be considered further in chapter 6.

For our present purposes the question is: Can the structure of a sentence in one language influence the production of structure in another language? The answer appears to be yes, although with some caveats. Several studies have now found priming between languages with a number of different structures (Bernolet, Hartsuiker, & Pickering, 2007; Hartsuiker, Pickering, & Veltkamp, 2004; Loebell & Bock, 2003; Schoonbaert, Hartsuiker, & Pickering, 2007). Surface similarity between the prime sentence and the target sentence appears to be important for bilingual priming, but this is not unique to bilingual representations and it appears that in general structural priming is sensitive to the degree of similarity or overlap of structural form (Bock & Loebell, 1990; Pickering, Branigan, & McLean, 2002).

Loebell and Bock (2003) examined syntactic priming in German–English bilingual speakers. Participants in their study first heard and then repeated a sentence in either German (their first language) or English (their second language) and then described an unrelated picture in the other language. While the content of the picture did not share anything in common with the repeated (prime) sentence, it could best be described using a dative construction in either language. Dative constructions in English and German can come in one of two forms, which we can label the double object and the prepositional object. The double object version of the dative construction looks like this: *The boy gave the man a book*, whereas the prepositional object looks like this: *The boy gave a book to the man*. Notice that the order of the two nouns following the verb changed, and that in one case (the prepositional object version) one of the nouns is put into a prepositional phrase. Loebell and Bock (2003) found that the production of German datives by their participants primed the production of English datives, and vice versa.

In another study, Hartsuiker et al. (2004) looked at syntactic priming in Spanish–English bilinguals. They found that participants in their study were more likely to produce English passive sentences when they had just encountered a Spanish passive, compared with hearing either an active or intransitive sentence. They account for their results using a lexically driven model in which combinatorial information about what kind of argument a verb can appear with is stored as part of the lexical information about the verb—at the lexical level (Pickering & Branigan, 1998). Such information is stored, roughly speaking, in "combinatorial nodes." For example, there could be a "passive" node, and all verbs that could appear in the passive voice would be linked to this node. Hartsuiker et al. extended this model to bilinguals by having a shared lexicon with each word tagged for language. This means that verbs from both languages could potentially be linked to the same combinatorial node.

This lexical approach is supported by another influence on priming—the "lexical boost." While lexical repetition is not necessary to priming between sentences and languages, it appears to help. Several studies have shown that priming is enhanced when the verb is repeated between prime and target sentence in monolinguals (Branigan, Pickering, & Cleland, 2000; Corley & Scheepers, 2002; Pickering & Branigan, 1998). In a cross-language priming context, verbs are not repeated, but the translation of the verb in the prime sentence is used in the target utterance. So, if the verb in the target sentence is "offers" then the prime verb would be the translation of *offer* in the prime sentence language that's being tested. Schoonbaert et al. (2007) examined this boost effect in Dutch–English bilinguals by conducting four experiments that tested both within language and across-language priming in this way. They found the same boost effect for within-language priming in both English and in Dutch for their bilingual participants, but only found a boost effect in across-language priming when the priming was from Dutch to English—from the first language to the second language. There was priming from English to Dutch, but no additional priming when the verb used was a direct translation between prime and target. They largely adopt Hartsuiker et al.'s (2004) model, but propose that lexical processing needs to be more carefully considered as it is most likely the underlying cause of this asymmetry, with weaker lexical–conceptual links for the L2 compared with the L1.

One caveat on the findings thus far in this area is that we are not yet certain to what extent proficiency influences the degree of priming. Also, the languages tested to date are all reasonably similar to each other, and so we do not know whether structurally dissimilar languages (which might have occasional overlaps to test) would show such priming. At this point, the results allow us to believe that in reasonably proficient bilinguals of structurally similar languages, there is a certain degree of interaction and possibly shared representations between those languages.

THE BROADER IMPACT OF HAVING MORE THAN ONE LANGUAGE

Another big issue when considering multiple language representation is how having more than one language might impact language processing, and whether it will have any influence on how other, nonlanguage skills function. We saw in the previous sections that languages may interact in interesting ways, but we haven't really seen what the larger impact of speaking more than one language might be. For example, does speaking more than one language influence how quickly you can access words in your mental dictionary? What about nonlanguage tasks—does speaking two languages allow you to switch between tasks more easily than someone who speaks only one language? Another hot topic in language processing is looking at the impact of bilingualism more generally.

Bialystok (2008) provides a nice synthesis of the research in this area, and the results are mixed. In terms of their mental dictionaries, there appears to be actually a small disadvantage for fluent bilinguals related to the ease with which they can access words. For example, you have almost certainly experienced the feeling that a word is on the tip of your tongue—you know that you know it, but you just can't quite get it out. This is actually called a "tip-of-the-tongue" (TOT) state by researchers and it provides a window into the word retrieval process. It turns out that bilingual speakers experience more TOT states than monolingual speakers (Gollan & Acenas, 2004). Other studies have also shown that bilingual speakers take a little longer on average when asked to name pictures of objects (e.g., Roberts, Garcia, Desrochers, & Hernandez, 2002). It's not entirely clear why bilinguals have more difficulty with lexical access than monolingual speakers, although one possibility is that bilingual speakers have somewhat weaker links between the connections necessary for word retrieval (Michael & Gollan, 2005). To see why this could cause problems, think about the difference in language use between bilingual speakers and monolingual speakers. If a monolingual speaker of English wants to refer to a dog, then they pretty much have the one word that they will use, *dog*. However, a bilingual speaker of English and Spanish might use either *dog* or *perro*, depending on the context. This means that, on average, the frequency of use for *dog* will be higher for monolinguals than the frequency of use for bilinguals. The fact that *dog* has different frequencies for these different speakers may have an impact on the strength of the representation of that word in the lexicon between the speakers. Bilinguals may have somewhat weaker representations because they simply use words in a given language less frequently than speakers who only have one language to use.

There are advantages to being bilingual, too (aside from the obvious one of being able to communicate with a bigger set of people!). Bilinguals perform better than monolinguals in tasks involving "executive control" (e.g., Bialystok, 2006; Costa, Hernandez, & Sebastian-Galles, 2008). Executive control is a bit of an umbrella term that encompasses several functions, including the ability to shift tasks, update information in working memory, and control attention. In essence, bilingual children and adults appear to be better at tasks that test these skills than monolingual speakers. One reason for this may be that bilingual speakers have a great deal of experience managing attention and task shifting between two or more systems—namely their languages. This experience then contributes to a greater ability to do this in nonlanguage tasks as well (Bialystok, 2008). These results also suggest an important and nontrivial connection between language representation, processing, and other cognitive systems.

WRAPPING UP

In this chapter, we have covered several key topics related to how languages are represented psychologically and neurologically, focusing primarily on bilingual representations. The results from many areas of research show that multiple languages may be represented in the same areas of the brain, but that the degree of overlap may be due in part to one's skill in the language. Bilingual representations of language interact during language processing and appear to have impacts even beyond language use.

REFERENCES

Baddeley, A. (1992). Working memory. *Science, 255*, 556–559.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 931–949.

Bialystok, E. (2008). Bilingualism: The good, the bad and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3–11.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology, 60*, 68–79.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35*, 1–39.

Branigan, H., Pickering, M., & Cleland, A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*(2), 13–25.

Caramazza, A., Capitani, E., Rey, A., & Berndt, R. (2000). Agrammatic Broca's aphasia is not associated with a single pattern of comprehension performance. *Brain and Language, 76*, 158–184.

Caramazza, A., & Costa, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition, 2*(3), 231–244.

Caramazza, A., & Zurif, E. (1976). Dissociations of algorithmic and heuristic processes in sentence comprehension: Evidence from aphasia. *Brain and Language, 3*, 572–582.

Chee, M. W., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W., Thiel, T., et al. (1999). Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron, 23*, 127–137.

Corley, M., & Scheepers, C. (2002). Syntactic priming in English sentence production: Categorical and latency evidence from an internet-based study. *Psychonomic Bulletin & Review, 9*, 126–131.

Costa, A., Albareda, B., & Santesteban, M. (2008). Assessing the presence of lexical competition across languages: Evidence from the Stroop task. *Bilingualism: Language and Cognition, 11*(1), 121–131.

Costa, A., Hernandez, M., & Sebastian-Galles, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*, 59–86.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., et al. (1997). Anatomical variability in the cortical representation of first and second language. *NeuroReport, 8*, 3809–3815.

Dickey, M. W., & Thompson, C. (2009). Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eye-tracking. *Journal of Neurolinguistics, 22*, 563–583.

Drai, D., & Grodzinksy, Y. (2006). A new empirical angle on the variability debate: Quantitative neurosyntactic analyses of a large data set from Broca's aphasia. *Brain and Language, 96*, 117–128.

Galton, C. J., Patterson, K., Graham, K., Lambon-Ralph, M. A., Williams, G., Antoun, N., et al. (2001). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology, 57*, 216–225.

Gollan, T. H., & Acenas, L. A. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish–English and Tagalog–English bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition, 30*, 246–269.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish–English bilinguals. *Psychological Science, 15*(6), 409–414.

Hermans, D. (2004). Between-language identity effects in picture–word interference tasks: A challenge for language-nonspecific or language-specific models of lexical access? *International Journal of Bilingualism, 8*(2), 115–125.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition, 1*(3), 213–229.

Illes, J., Francis, W. S., Desmond, J., Gabrielli, J. D. E., Glover, G., Poldrack, R., et al. (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain and Language, 70*, 347–363.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580–602.

Linebarger, M., Schwartz, M., & Saffran, E. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition, 13*, 361–392.

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics, 41*, 791–824.

Michael, E., & Gollan, T. H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. F. Kroll & A. M. B. de Groot (Eds.), *The handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). New York: Oxford University Press.

Nicol, J. (Ed.). (2001). *One mind, two languages*. Malden, MA: Blackwell.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain—Proficiency and age of acquisition of the second language. *Brain, 121*, 1841–1852.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*, 633–651.

Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent structure is formulated in one stage. *Journal of Memory and Language, 46*, 586–605.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*, 576–582.

Reilly, J., & Peelle, J. E. (2008). Effects of semantic impairment on language processing in semantic dementia. *Seminars in Speech and Language, 29*, 32–43.

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology, 16*, 635–645.

Roelofs, A. (2003). Shared phonological encoding processes and representations of languages in bilingual speakers. *Language and Cognitive Processes, 18*(2), 175–204.

Roland, D., O'Meara, C., Yun, H., & Mauner, G. (2008). *Discourse and object relative clauses: The effect of topic versus mention*. CUNY 2008 Conference on Human Sentence Processing. University of North Carolina, Chapel Hill, NC.

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals:

Evidence from syntactic priming. *Journal of Memory and Language, 56*, 153–171.

Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: A form of circumscribed cerebral atrophy. *Behavioral Neurology, 2*, 167–182.

6

Language in the Real World: Dialogue and (Co)reference

Up to now we have largely considered how language comprehension and production work in isolation—how do readers process information from language, deal with ambiguity, and so on. But, most daily language involves interaction with other people—language occurs in a conversation or dialogue. In this chapter, we will turn to questions related to how speakers and hearers influence each other. A major focus in this area is on how speakers are influenced by their audience, either by what someone has just said to them or by what they know (or think they know) about what their listener knows. But, the listener, too, can be influenced by what they think about the speaker. And so, in this chapter, we will first look at research on dialogue, and especially how a dialogue context influences speakers (and in some cases, listeners). Then, we will look at a specific aspect of linguistic form—how we choose to refer to something, and what impact this has on readers and listeners.

RESEARCH ON DIALOGUE

Speakers clearly have an impact on their listeners; after all, the whole point of speaking is to transfer some kind of information from speaker to hearer, but what about the other way around? Do listeners also have an impact on speakers? And, if so, what kind of impact? One thing to keep in mind, of course, is that in a typical dialogue the speaker and listener take turns, and so the boundary between speaker and listener is thus somewhat fluid. The term "interlocutor" is often used generally to participants in a dialogue, without specifying whether they are currently the speaker or listener.

The usual goal of a dialogue or conversation is successful communication (of ideas, information, intentions, and so on), and so it would make sense that a speaker would pay careful attention to the needs of a listener and do things like avoid ambiguity and package information in a way that flags particular information as important or new to the listener. In fact, speakers may well do these things, as we shall see, but they may not do them specifically to make their utterances easier for their listeners. Despite the ease with which we appear to do it, language production is not a simple thing and so there may be things that speakers (unconsciously) do to make their own production easier and more fluent. Also, as we shall see when we discuss one model of dialogue, there appears to be a systematic push in dialogue for speakers and listeners to influence each other's language representations at every level so that over time, interlocutors sound more and more alike.

It is also important at this point to note that the way that experimenters use the term "dialogue" may not exactly match what you normally think of. The contexts that are used in most experiments are necessarily fairly constraining, in order to be able to control for extraneous factors that are not of interest in these particular experiments. Also, and again this is related to control, in many cases the other person in the dialogue isn't even a real participant in the study, but instead is a "confederate" of the experimenters. This may initially seem suspicious because perhaps the lack of a "real" participant in the dialogue influences its results. In fact, there is some debate about the use of confederates among researchers, but in studies that have used confederates, participants are asked at the end of the experiment if they think their partner was a real or fake participant and they rarely report that they think their partner isn't real, even when they are given financial incentives to guess correctly about it (Keysar, Barr, Balin, & Brauner, 2000). Further, on the rare occasions when the real participants do report suspecting their partners, experimenters do not analyze their data, or analyze it separately. In studies where it was analyzed separately, there did not appear to be a difference from participants who were oblivious. Interestingly, the participants' perceived role in a dialogue does seem to matter. In one study, Branigan, Pickering, McLean, and Cleland (2007) found that it mattered whether people considered themselves to be a direct participant in a dialogue or a bystander.

TAKING THE LISTENER INTO ACCOUNT DURING DIALOGUE

One of the key things to keep in mind when looking at dialogue is that it is at heart a joint activity between speakers in which they must cooperate in order to understand the dialogue (Clark, 1996). As mentioned earlier, there is not a static speaker and a listener, but instead there is a dynamic situation in which both speaking and listening are coordinated, with roles changing frequently such that listener needs to listen with the possibility that he or she will soon be the speaker. The idea of coordination is not unique to dialogue—there are many activities that we engage in that are joint in nature, for example ballroom dancing and sawing using a two-handed saw are activities that are commonly pointed to as joint activities akin to dialogue. In such activities, both participants must work together to be successful. A key question for interlocutors in a dialogue is how they successfully coordinate their joint language activity. For psycholinguistics, it is of particular interest what kinds of information are used, and when, in order to make this work in real time. There are a number of possibilities, actually, which we shall now turn to in the next section.

COMMON GROUND AND THE EGOCENTRIC SPEAKER

A key idea in thinking about dialogue is the notion of common ground (Stalnaker, 1978). Common ground is what interlocutors can reasonably assume is shared knowledge, given the evidence. For example, common ground may include knowledge about the town or city one lives in. Consider how you might describe how to get to a local coffee shop to someone you know lives just down the street from you, compared with someone who is visiting for the first time from out of town. Or, consider how you might tell a sibling about a recent big event in the family (like a wedding) compared with a colleague or coworker who you do not know well. In these cases, you share different degrees of common ground on these topics, and the idea is that this should influence the dialogue. Common ground can also be the immediate context, linguistic or nonlinguistic. For example, if you have just told someone that you have a sister, you can then reasonably assume that this sister is now in common ground between you and your interlocutor, and may be so in any future conversations.

On one hand, there is no question that common ground plays a role in dialogue, and there is evidence that it influences the use of referential expressions, and in particular seems to help guide how we refer to things (e.g., Clark, Schreuder, & Buttrick, 1983; Clark & Wilkes-Gibbs, 1986). It forms a key part of the notion of "audience design" in which the speaker draws from the mutual knowledge that he or she shares with his/her interlocutors.

However, from a computational perspective, common ground is costly: there are a huge number of potential things that could be in common ground between interlocutors, even relative strangers. Further, common ground requires keeping distinct what you know from what you can reasonably infer that your interlocutors know. And this information needs to be updated in real time. Many researchers have argued that this cost is too high to expect interlocutors to keep and monitor this type of common ground and have proposed alternative models for how something like common ground might work. Thus, while common ground is clearly a part of language in dialogue, the question is how it is used. Horton and Keysar (1996) examined whether common ground is involved in the initial stages of utterance planning, or if it is instead used during a later stage of processing in which speakers monitor their utterances to detect (and hopefully intercept) any deviations from the message they intend to convey, as well as any violations of well-formedness (i.e., speech errors). To test these two possibilities, Horton and Keysar used a modified referential communication task in which participants would see one-half of a computer screen and another person would see the other. Participants would see two objects on their side of the screen, one over the other, and then the object at the top of the screen would move to the other person's side. The participants were asked to describe the moving object in a way that would allow the other person to identify whether the object on their side of the screen was the same as the one that was just on the participants' side of the screen. Half of the participants were in a "shared context" version of the study, in which they knew that their listener had a copy of the lower object on their side of the screen, too. So, in this case, both speaker and listener might see, for example, a large circle at the bottom on their displays. The other half of the participants were in a "privileged context" condition, in which speakers knew that the listener did not have any other object shown on their side of the screen. So, in one case the speakers knew that they had common ground with the listener, and in the other they knew they did not. The question was whether the speakers would use this knowledge or not. For example, if the speaker saw a large circle at the bottom and smaller circle on top, he or she might say "a small circle" to distinguish it from the bigger circle, if he or she knew the listener also saw the bigger circle. However, it would be faster to simply say "a circle" if there was no other shape in common ground. Horton and Keysar found that when under no time pressure to produce an utterance, speakers used an expression that was appropriate for their context: speakers in the privileged condition were more likely to produce a shape name with no modifier (circle), whereas speakers in the share context condition were more likely to produce names with a modifier (small circle). This suggests that speakers were indeed taking common ground into account in this task. However, when the speakers in the experiment were placed under time pressure by being told they had to begin their description as soon as the object on top began to move (1.5 s after it had appeared on screen), they then produced adjectives roughly half the time, regardless of whether they were in the privileged or shared condition. This suggests that while common ground in used, it is more likely part of the later, monitoring process that speakers use rather used during the initial stages of production.

Looking at comprehenders for a moment, it appears that they, too, rely on common ground, although also not initially. Keysar et al. (2000) used a referential communication game in which a real participant and a confederate participant sat facing each other with a vertical array of objects between them.

These objects were placed in little cubbies, and in some of the cubbies both sides were open so that both participants could see the object, but in others one side would be blocked, so that only the real participant could see the object. In this experiment, the confederate was the assigned to the "director" role and gave commands like "Pick up the small candle" to the real participant—the "addressee"—who would then comply with the commands. Because of the blocked cubbies, some objects were not in common ground and, thus as far as the addressee was concerned, were not known to the director.

The addressee had their eye gaze monitored while they participated in the task. This is because eye gaze is known to be connected in real time with processing (e.g., Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995) such that when people are listening to speech and looking at a visual scene, they will look at the objects that correspond to what they are hearing. For example, as we saw in chapter 3, if someone hears "Pick up the candle" while looking at an array of objects that includes a candle, then they will look at the candle as soon as they can identify it as the referent of what they heard. Importantly, if there is more than one candle in the scene, then they will look at both candles until one can be distinguished. While this type of situation can lead to delays in identifying the correct object of reference, it can also lead to faster identification times: if there is a small candle and a large candle and the listener hears "Pick up the small candle" then their eye gaze and will shift to the smaller candle even before the word "candle" has been fully presented (Sedivy, Tanenhaus, Chambers, & Carlson, 1999). So, in Keysar et al. (2000), the key question was whether the addressee's eye gaze would show that they were considering the blocked objects as potential referents, even though they were not in common ground and could not be referred to by the director. If they ignored objects that were not shared, then this would show that information about common ground was used during initial processing; however, if addressees looked at objects not in common ground then this would suggest that they were not consulting common ground information, at least during the earliest part of reference determination. The results from Keysar et al. show that listeners, too, adopt an "egocentric" perspective when initially processing language—the results from the eye gaze data clearly showed that the addressees did not restrict their visual search for referents to only those object that could be seen by both participants. Taken together with the results from speakers, it appears that common ground is not consulted during initial stages of production or comprehension, although it is used later on.

INTELLIGIBILITY

When we speak we can pronounce the same word in different ways: we can pronounce a word carefully so that it is maximally intelligible, or we can pronounce it in a more "sloppy" manner, perhaps skipping quickly over certain, unstressed sounds, or reducing vowels to a middle "uh" type sound. We can, of course, choose to do this deliberately, but what about when we're not thinking about it? Some research has shown that when we say a content word (like a noun or verb) for the first time in a larger discourse we say it with higher intelligibility than when we say it again, meaning that it is more recognizable when it is heard out of context (Fowler & Housum, 1987). This makes sense from a communicative perspective because as a new word it may be harder to determine by the listener just from context compared with something that is previously given, either by explicit mention or by implication. Related to this, some work has shown that it is not simply repetition that causes this effect because using the same term to refer to a new object causes an increase in intelligibility (Bard, Lowe, & Altmann, 1989).

As we saw in the previous section, there are two possibilities: one is that speakers are consistently tailoring their intelligibility for the benefit of their listener, making new or unpredictable words clearer. However, as already mentioned, this is a somewhat costly process in terms of processing because it will mean constantly monitoring the availability and givenness of particular referents for the listener. The other possibility is that speakers initially pay attention to their own knowledge about the state of the discourse and world knowledge, and the intelligibility of their speech reflects this view. In most cases in the real world, a speaker's own knowledge about the current discourse will be in sync with their listener and so it will not make a difference whether the speaker is consulting his/her own knowledge or making a calculation about their listener's knowledge. However, in cases where there is a mismatch between speaker and listener, this strategy could backfire: the speaker could refer to something unknown to the listener as if it was known. In this case, the speaker would need to adjust his or her speech after noticing the problem, either from monitoring his/her own speech or receiving feedback from the listener indicating that they didn't understand.

Bard et al. (2000) looked at the effect of having a listener on the intelligibility of speakers' utterances in over 100 unscripted conversations that involved a mapping task in which two participants worked together to reproduce one participant's route on a map onto another map. Crucially, participants could not see each others' maps, and the maps were not entirely identical. Also, participants swapped partners during their experimental session, so that a speaker would have two listeners over the course of the experiment. This creates an ideal circumstance to see whether speakers adjust their speech to a new listener or not. Bard et al. compared how intelligible a speaker's first mention of a landmark was with their first listener to how intelligible their first mention of a landmark was with their second listener. In the first case, the mention is new for both speaker and listener, but in the second case it is only new for the listener—the speaker has already mentioned and is familiar with the landmarks on the map. The researchers found that speakers' intelligibility was significantly lower for their second listener compared with their first, suggesting that speakers are not "resetting" their utterances for a new listener. This supports the idea that speakers are not necessarily carefully monitoring their own utterances with respect to common or shared information.

AMBIGUITY AVOIDANCE

In chapter 4, we saw that language has the potential to often be ambiguous both in terms of individual word meaning as well as structure. Yet, such ambiguity may be avoided depending on the speaker's choice of words and so a natural question is whether, and when, speakers appear to avoid ambiguous language. A lot of work on this topic has been done with speakers in isolation (Arnold, Wasow, Asudeh, & Alrenga, 2004; Ferreira & Dell, 2000; Kraljic & Brennan, 2005) and in general has found that speakers may produce more or less ambiguous structures, but do not do so based specifically on avoiding ambiguity. Haywood, Pickering, and Branigan (2005) looked at whether speakers avoid ambiguity during dialogue using a communication game in which participants took turns giving and receiving instructions about how to move objects around in an array. One of the participants was a confederate, but from the real participants' perspective, the experiment involved a game in which he or she would sit in front of a set of objects and geometrical shapes that were arrayed in a 5 × 5 grid. He or she would pick up a card that pictured the same array, but with one object in a new place. The goal of the participant was to tell the other participant what to do to make the array of real objects look like the picture (without showing the picture). Then, the other participant would move the object, and then pick his or her own card and do the same thing. Thus, the participant took turns both giving instructions and receiving them. However, the confederate's speech was scripted to be either ambiguous (based on context) or unambiguous. For example, the set of objects in the grid might include both a small stuffed pig and another small stuffed pig that is sitting on top of the block. In this case, if the speaker says "Put the pig on the block..." it is ambiguous whether "on the block" is specifying which pig or where to put the pig. On the other hand, if the speaker says "Put the pig that's on the block..." then this is no longer ambiguous. If there is only one pig, then either instruction is unambiguous—the ambiguity thus depends on both the syntactic form and the context. The question was whether the real participants would avoid ambiguous instructions when the context required it, or not. Another possibility is that the participants' instructions would simply echo the syntactic structure of what they just heard (recall syntactic priming from chapter 5), regardless of whether the context was ambiguous. The experimenters also varied whether the confederate was "helpful" (avoided ambiguous themselves) or "unhelpful" (used ambiguous instructions). Haywood et al. found that participants produced *that's* in only 32% of their utterances, and that they produced more when the confederate had just produced one. However, they also produced more *that's* when the context meant that their instruction would be ambiguous without it. This effect, while small (only an 8% difference) suggests that speakers can take ambiguous contexts into account and produce sentences that avoid the ambiguity. This effect is unconscious—post-experiment interviews indicated that none of the participants were aware of the syntactic ambiguity, nor the confederate's helpfulness (or lack thereof).

PRIMING AND ALIGNMENT: A MODEL OF DIALOGUE

It appears that speakers do not always or initially consider all the needs of their speakers. However, even setting aside the rather obvious point that we respond to the content of what we hear, it is also clearly the case that speakers and listeners do influence each other during dialogue. In fact, there appear to be a number of low level influences, particularly due to priming, that appear to cause interlocutors to converge on similar speech over the course a dialogue. Pickering and Garrod (2004) argue that various forms of priming, including syntactic priming, serve to align different forms of representation and that in successful dialogue it is ultimately the mental states of the interlocutors that become aligned. This proposal suggests that during dialog, interlocutors are slowly bringing phonological, lexical, and syntactic representations into sync. In observable terms, this means that interlocutors come to produce words in the same way, use the same terms to refer to the same things, and even use the same syntactic structures, regardless of their content.

In terms of pronunciation, we have already seen that speakers reduce articulation and intelligibility over the course of a dialogue (e.g., Bard et al., 2000; Fowler & Housum, 1987). Interestingly, Bard et al. also found that this reduction occurred even when the speaker had only heard the word spoken by a different speaker. That is, this reduction occurs when the speaker either has previously said the word, or heard it spoken. Pickering and Garrod also point out that previous work has shown that interlocutors alter (and align) accent and speech rate over a discourse (Giles, Coupland, & Coupland, 1991; Pardo, 2006). Taken together, these studies lend support to the proposal that interlocutors are aligning the way in which they pronounce words.

There is also evidence in support of the idea of lexical alignment between interlocutors—that interlocutors converge on shared terms to refer to things when these things are referred to repeatedly (Brennan & Clark, 1996; Garrod & Anderson, 1987; Wilkes-Gibbs & Clark, 1992). Brennan and Clark (1996) investigated this phenomenon, sometimes called *lexical entrainment*, by having pairs of participants do a matching task in which one person would be the director and the other would be the matcher. Each participant was given an identical set of 12 cards with pictures on them of things like shoe, a car, a dog, and a fish. The precise set of pictures was manipulated experimentally, such that in certain trials there might be cards that depicted additional types of the same object, so that for example, there might be two cards with different types of dogs or two cards with different types of shoes. These cards were laid out in front of each of the participants in an array, separated by a partition so that the participants could not see each other's cards. They then needed to get the array of the matcher's cards to match the director's array. Because of the presence of different types of the same thing (e.g., multiple types of fish), participants would need to say more than simply "fish" to uniquely identify each fish. Brennan and Clark found, over a series of three experiments, that speakers did indeed converge on shared terms to refer to the objects on the cards. They argued in favor of the idea that interlocutors are in effect forming a conceptual pact in this process, in which a temporary agreement is made about how to conceptualize the object and thus refer to it.

Finally, with respect to syntactic alignment, there is in fact a great deal of interest and research on syntactic or structure priming. Priming in this case refers to an interesting phenomenon in speech production (both written and spoken): language producers tend to repeat structures that they have recently heard, read, or produced. For example, Bock (1986) found that when participants heard and then repeated sentences like "The referee was punched by one of the fans" they were more likely to use the same structure (a passive) when describing a following picture, even when the picture was not related in meaning to the previous sentence in any way. Following Bock (1986), more than 100 studies have used or tested syntactic priming in some way, including examining shared structural representations between languages in bilinguals, the relationship between speaking and writing, and the interaction of lexical and structural information (see Pickering & Ferreira, 2008, for an overview of research involving structural priming). Crucially for our purposes here, syntactic priming has also been examined in the context of dialogue, particularly because researchers have found priming from comprehension to production (e.g., Bock, Dell, Chang, & Onishi, 2007). This finding introduces the idea that when someone hears another person use a particular structure it makes that person more likely to use that same structure in turn. This same finding has been found in a number of studies involving a dialogue situation (Branigan, Pickering, & Cleland, 2000; Cleland & Pickering, 2003; Haywood et al., 2005). For example, Branigan et al. (2000) found that speakers were more likely to describe the picture on a card as "The cricketer giving the plate to the diver"—in which the verb *giving* is followed by an object *the plate* and then a preposition *to the diver*—after they had just heard another use the same type of structure.

Taken together, these studies suggest that interlocutors are indeed converging on similar ways to convey information, in effect aligning representations at the lexical and syntactic levels. Further work will help elucidate how this process works, including to what extent common ground is consulted in this process and whether lexical information is needed to support syntactic priming. Most importantly, research in this area has established that interlocutors have a measurable effect on one another at a number of levels, and that this unconscious alignment of internal representations means that speakers and hearers may not need to use common ground information in order to communicate effectively.

COREFERENCE

Another area where coordination between speaker and hearer can be critical is in how speakers (or writers) make reference to things or events, either for the first time or to refer to something already mentioned (called coreference). The words they use to do this are called coreferring expressions. For example, pronouns are coreferring expressions *extraordinaire*—they refer to and take their meaning from things previously mentioned or already known.

But, of course, pronouns are not the only kind of coreferring expression—there are others forms of reference, including category noun phrases such as *the bird*, which could refer to any type of creature that both speaker and hearer acknowledge as belonging to this category. In order to understand how coreferential processes work, we must first think about exactly *what* pronouns and other forms of coreference are referring to. One way to think about how discourse works is that each interlocutor is building a mental model of the situation under discussion (e.g., Johnson-Laird, 1983). This mental model is a representation of the things, relationships, and attributes involved in the dialogue. This is not a static model by any means, and changes with each new utterance. Parts of the mental model might be temporarily more accessible than others, depending on what had just been mentioned or even seen by the interlocutors. So, we have a discourse on one hand, with (co)referring expressions (i.e., words) and a mental representation of the meaning of the discourse on the other, with referents. This allows for a subtle but important distinction to be made. Let's consider the following example:

1. John went to the concert. Jill wanted to go with him.

The *him* here has very little meaning of its own and its interpretation depends on the context, in this case the previous sentence. Both the linguistic context and the mental model that comes from it are important to interpretation: *him* is interpreted in terms of the word it links with (*John*, in this case) as well as the representation of John in the mental model of the discourse (the representation of John). The word itself that a pronoun links to is some times called the antecedent or antecedent expression, and the representation in the mental model is called the antecedent referent. As we shall see, this dual linking really does matter in how we interpret coreferring expressions, especially pronouns.

The status of the referent in a mental model matters to how a speaker or writer might refer to that referent: When speakers refer to something, they have a number of options in terms of *how* they refer to it. For example, here is a quick discourse:

> Let me tell you about my cat, Dmitri. My husband and I found Dmitri at our local shelter when he was a kitten and fell in love with him right away, despite the fact that he was the only one in the shelter who was hidden under the newspaper in his cage, howling pitifully.

In this discourse were have a bunch of (co)referring expressions and corresponding referents. Let's focus in on the cat, Dmitri, and pay close attention to how I referred to him: notice that I first described him with a category name (cat) and then gave his proper name (Dmitri). After that, I used his proper name again, but then switched to a pronoun (*he*) for the rest of my references to him (until I changed the subject and began to discuss how I wrote about him). The key thing to notice here is that I have used a number of different forms of coreferring expression to refer to the same referent—my cat—over the course of a very short passage about him. This is a critical part of how we use language in the real world—as speakers and writers, we use different forms of expression to refer to the same thing. As comprehenders, we use these forms as cues to the status of the things the speaker is talking about. Thus, the choice of the form of referring expression is not entirely due to the whim of the speaker, but is dependent in part on the structure of the discourse and the status of the things in it. This becomes clearer when considering the same passage but with a few words changed:

> Let me tell you about my cat, Dmitri. My husband and I found Dmitri at our local shelter when Dmitri was a kitten and fell in love with Dmitri right away, despite the fact that Dmitri was the only one in the shelter who was hidden under the newspaper in Dmitri's cage, howling pitifully.

But, then, only using pronouns doesn't work, either:

> For example, let me tell you about him, him. My husband and I found him at our local shelter when he was a kitten and fell in love with him right away, despite the fact that he was the only one in the shelter who was hidden under the newspaper in his cage, howling pitifully.

Neither of these passages sounds quite right. In the first case, you may keep wanting to think there is, in fact, another Dmitri that is being spoken about. In the second, one must assume that I have previously introduced the "he" of which I am writing. So, in general, we use descriptions like categories (the cat) to introduce new things, but then must switch to some alternate way of referring to them, like pronouns. Some languages, like Spanish and Chinese, don't even need to use pronouns overtly, but instead allow the reference to go completely unsaid. It is *understood* that the speaker is referring to some previously mentioned and familiar thing. Notice that pronouns have less information built into them—in English they supply gender and number and not much else. In a language like Chinese, there may be nothing at all. Thus, overt pronouns in English and "null" pronouns in Chinese are considered to be *reduced* forms of referring expressions. In general, then, we can say that speakers tend to use fuller forms of expression to refer to new things and reduced forms to refer to known things.

But, often lots of things are known, and lots of reduced forms thus get used. Because pronouns do not provide a lot of information about what they refer to, they provide yet another source of ambiguity in language. Consider:

Joe hit John and then he kicked him.

Who do the *he* and *him* refer to? After reading this, most English speakers will say that *he* refers to Joe and *him* refers to John. Indeed, outside the context of a discussion of pronouns, people may not even notice that another interpretation is possible, namely one in which John retaliated from being hit by in turn kicking Joe. This interpretation is easiest if we stress the pronouns:

Joe hit John and then HE kicked HIM.

So, we have a strong preference for one interpretation, although there is technically room for ambiguity. In fact, people generally have no problems understanding what particular pronouns refer to, even the unspoken ones used in languages like Chinese. Yet, if one goes to write an algorithm that would allow computers to understand the references of pronouns in human speech, things get tricky. There is no single, clear cue or pattern that applies. And so, like so much of language, it is effortless for us to do but much more difficult to understand how we do it.

The form of referring expression is very important for a successful dialogue, after all, we want our listeners to understand what we are referring to. As we discussed earlier above, interlocutors appear to align over time in terms of the labels that they give in referring to novel objects. But, what about when reference shifts to the reduced forms we just discussed? For example, how do listeners or readers determine the reference of a pronoun? As already hinted, there is no straightforward answer to this, but instead there are a number of cues that appear to be relevant.

In the rest of this chapter we will discuss some of the key factors that have been found, divided into two categories, based on the two ways in which pronouns are interpreted that were mentioned earlier—the antecedent expression and the antecedent referent. First, from the link to the antecedent expression, there are influences based on the language itself—constraints based on the syntactic rules of the language. Second, based on the link to the antecedent referent, there are influences based on the statuses of the possible referents in the mental model of the discourse, including the accessibility of the referents and what the relationships between them are. Of course, these statuses and relationships often come from the language of the discourse, but in this case the influence from the particular language used may be more indirect. For example, mentioning something first in a sentence may increase the accessibility of the referent in the mental model, which in turn may make it easier to interpret as the antecedent of an otherwise ambiguous pronoun. This is in contrast to a situation in which the syntactic position of the word itself influences coreference.

LANGUAGE-RELATED FACTORS

There are some constraints and preferences on how to interpret pronouns and other coreferring expressions that appear to be structural or syntactic in nature. For example, there appears to be a preference (all else being equal) for unstressed pronouns to refer to an antecedent that shares the same grammatical role. This is referred to as grammatical role parallelism (e.g., Stevenson, Nelson, & Stenning, 1995), and you can see it at work in the example earlier about Joe and John, repeated below:

Joe hit John and then he kicked him.

However, while this preference seems to be grammatical in nature, it appears to also depend on the semantic relationship between the two clauses and is most clear when there is a connective like "and" between them.

There also appears to be relatively hard and fast constraints on how to interpret pronouns: in particular, different types of coreferring expressions are argued to be subject to the principles of binding theory (Chomsky, 1981). Binding theory is intended to account for why in (2) below, *himself* cannot refer to John, even though John is the only possible referent in the sentence.

2. John said that Mary hurt himself.

Principles of binding theory are based on a notion of "c-command," which has to do with the hierarchical relationships between phrases in a sentence. Roughly speaking, a phrase X is said to *c-command* phrase Y if (and only if) every phrase in the sentence that contains X also contains Y. The name "binding theory" comes form the fact that when phrase Y is coreferential with phrase X and c-commanded by it, then Y is said to be "bound" by X. With respect to (2) earlier, Principle A states that a reflexive coreferring expression must be bound by an antecedent that is within the same clause. *John* and himself are not in the same clause and so *himself* cannot refer to John. Conversely, Principle B says that pronominal coreferring expressions must *not* be bound within the same clause, and so in (3), *him* cannot refer to Stan.

3. Mary said that Stan hurt him.

In this way, the form of the referring expression can impose constraints on possible antecedents. While it is certainly the case that grammatical constraints apply (it really is not possible to have *himself* refer to *John or him* or to refer to *Stan* in these examples), the question is how these constraints apply during processing. One approach has proposed that grammatical constraints apply from the very beginning of processing and in fact serve as an initial filter on which antecedents might be considered from the very beginning (e.g., Nicol & Swinney, 1989). Another possibility is that this constraint applies when selecting among possible antecedents. In our (2) example, the difference between these two approaches is that in the first case *John* is never considered as a possible antecedent of *himself*, whereas in the second approach *John* is initially considered but then rejected.

Notice in these examples that gender, too, plays an important role in interpretation—in (2) *himself* cannot refer to John because of binding, but it also cannot refer to Mary. Although Mary is grammatically possible in terms of structure, she is not compatible with the gender of the reflexive. Although English does not make use of grammatical gender (e.g., dividing nouns into categories like feminine or masculine), in languages that do, like French and Spanish, the grammatical gender of the noun also matters—if the grammatical gender of a noun is, for example, masculine, it cannot be the antecedent of a feminine pronoun. Even "stereotypical" gender has an effect on pronoun interpretation—many nouns, especially names of professions, have a gender that is typically associated with them. For example, truck drivers are typically thought of as male, although there can be (and are) female truck drivers. Having a mismatch between stereotypical gender and pronominal gender can cause longer processing times (e.g., referring to a truck driver as *she*). Gender is a major constraint on pronoun interpretation and many studies have taken advantage of it in order to investigate aspects of pronoun processing, including how and when binding applies to pronoun resolution.

In a series of experiments, Badecker and Straub (2002) used gender to test the constraints of binding on interpretation. They found that varying the gender of a noun phrase that was inaccessible according to binding still had an effect on reading times after the pronoun. This is intriguing because if the noun phrase was truly inaccessible (i.e., not under consideration as an antecedent), then we would not expect it to have an influence on processing the pronoun. This issue was also investigated by Sturt (2003), who used eye-tracking to investigate the processing of reflexive pronouns in short passages like those given in (4) and (5) below.

4. Jonathan was pretty worried at the City Hospital. He remembered that the surgeon had pricked herself/himself with a used syringe needle. There should be an investigation soon.
5. Jennifer was pretty worried at the City Hospital. She remembered that the surgeon had pricked herself/himself with a used syringe needle. There should be an investigation soon.

In these passages, there was a prominent discourse character (e.g., Jonathon/Jennifer) who either matched or mismatched in gender with a reflexive pronoun. However, this character was syntactically inaccessible—it could not grammatically be the antecedent of the pronoun. In addition, there was a second character (e.g., surgeon) that was accessible to the pronoun, but matched or mismatched in stereotypical gender with the pronoun. Sturt found that readers did immediately apply binding constraints—while the stereotypical gender influenced the earliest measures of reading at the pronoun, the gender of the inaccessible character did not. However, Sturt also found that the gender of the inaccessible characters *did* have an influence on later processing measures in the sentence, including showing longer rereading times of the pronoun when the accessible character mismatched in stereotypical gender but the inaccessible character matched. Finally, a follow-up experiment looking at readers' final interpretation of these sentences revealed that they were more likely to (ungrammatically) interpret the named character as the antecedent of the reflexive when it matched in gender, and that this was especially true when the accessible antecedent (e.g., surgeon) mismatched in stereotypical gender. Taken together, these results led Sturt to propose that binding constraints form a "defeasible" filter in which binding constraints first apply but then may be overridden by other considerations. From these results we can see that grammatical constraints play an important role in processing pronouns.

FACTORS RELATED TO THE STATUS OF THE REFERENTS IN THE MENTAL MODEL

There are also a number of influences on coreference in general and pronouns in particular that appear to be best described as influences on the status of the referents in a representation (or mental model) of the discourse described in the discourse. In general, pronouns are used to refer to referents that are highly accessible in the mental model of the discourse (Ariel, 1990; Arnold, 1998; Chafe, 1976; Gundel, Hedberg, & Zacharski, 1993). There are number of ways in which a referent can be made more prominent or accessible in a discourse, resulting in faster processing times. Being previously and recently mentioned in the discourse (as opposed to unmentioned) is one of the main ways in which a referent may become more accessible, this is sometimes discussed as the givenness of referent (Gundel et al., 1993) and in fact, something need not be explicitly mentioned in order to be given. Cornish, Garnham, Cowles, Fossard, and Andre (2005) found that being a central part of an activity or state (e.g., when a woman is in the state of being pregnant, a baby is a central part) can be enough to make a referent accessible enough to cause no slowing in pronoun processing compared with an explicitly mentioned antecedent referent.

There are also different degrees of givenness and so different ways of being "given." First mention in a sentence appears to have an important impact, reflected not only in coreference processing, but also in other domains as well (Gernsbacher & Hargreaves, 1988). However, first mention is often confounded with being placed in subject position, and studies that separated these statuses have found conflicting results, with some studies finding that first mentioned referents cause faster processing than subject mentioned (Gernsbacher, Hargreaves, & Beeman, 1989) and other findings that subject mention causes faster processing (Cowles, Walenski, & Kluender, 2007). It may well be that both of these factors play a role in making referents more accessible.

An additional wrinkle in how to think about effects of first mention or subjecthood is that yet another status may be at work in these cases. For most simple, declarative sentences, one can partition the sentence into what it is "about" and what new information is being provided. What the sentence is about is the topic, and topics are often (although not always) at the beginning of the sentence. Some languages, like Japanese, overtly mark the topic with a special form, like a suffix. But, many languages do not do this and in such cases it can be difficult to separate effects of subject position, first mention, and topic. For our present purposes, it is enough to generalize to the following statement: except for specialized constructions, nouns that occur early in sentences are considered to be more prominent than nouns that occur later. Early mention in sentences often results in pronominal coreference later among speakers or writers (e.g., Cowles, 2007) and in faster reading times for pronouns that refer back among readers.

Speaking of production, some work looking at the production of coreferring expressions has found that the presence of alternate referents has an effect when pronouns are used. When there are two or more referents that are of the same gender, then it makes sense that speakers might avoid using an ambiguous pronoun. However, it appears there is more to it than ambiguity avoidance. Arnold and Griffin (2007) presented participants with two-panel cartoons depicting either a single character or two characters that were different genders. They also supplied participants with an initial sentence and then asked them to read that first sentence aloud and then produce another sentence that continued the story according to the second panel of the cartoon. They measured how often people used a pronoun to refer to the first-mentioned character in the given sentence and found that people used fewer pronominal references if there was a second character in the panels, even though in this case the character was unambiguously of the other gender. This suggests that not only ambiguity avoidance may be at work, but the presence of additional referents may affect the amount of attention that a speaker is able to give. Basically, if there is only one character in the scene, then the speaker can pay full attention to it, causing increased accessibility of that referent in the mental model of the scene. However, if there are two characters, then the speaker's attention is divided, and thus the accessibility (of both characters) is lower compared with when there is only one referent. In this account, pronoun use decreases because the referents are thus less accessible.

Word frequency appears to also play a role in how quickly people associate a coreferring expression with an antecedent referent. While some studies have not found significant changes in reading times for frequent versus infrequent antecedent expressions (e.g., Simner & Smyth, 1999), in an eye-tracking experiment, van Gompel and Majid (2004) found that using words that were less frequent to refer to antecedents caused *faster* reading times with pronominal coreference. They propose that encountering a less frequent word causes it to become more salient, which is supported by other work on word frequency showing that infrequent words are recognized better then frequent words (e.g., Glanzer & Adams, 1985; Malmberg & Nelson, 2003). This increase in antecedent saliency, which can be thought of as an increase in accessibility, causes faster reading times for coreferential pronouns.

Another potential source of influence on how people interpret pronouns comes from the notion of coherence relations. Sentences or sentence clauses in a coherent discourse can be said to have particular relations between them. For example, one sentence or clause may provide more information about an entity or event mentioned in the previous sentence, or *elaborate* that entity or event. Or, a second sentence may provide a cause for the state of affairs of the first sentence, or it may provide a consequence or result. It may provide an equivalent, or parallel relationship between other entities or events. Some recent work on coherence relations has provided evidence that they may, also, influence the interpretation of pronouns (Kehler, 2002; Kertz, Kehler, & Elman, 2006). For example, Kertz et al. (2006) compared parallel and result relations like the ones below:

6. Samuel threatened Justin with a knife, and Erin blindfolded him with a scarf.
7. Samuel threatened Justin with a knife, and Erin stopped him with pepper spray.

In (6), the coherence relation is parallel—Samuel does something and Erin does something, whereas in (7), the relation is one of result—Sam does something and as a result of this, Erin does something. Kertz et al. found that the type of coherence relation influenced how people interpreted the syntactically ambiguous *him*—parallel relations resulted in more interpretations of *him* referring to Justin, whereas result relations resulted in more interpretations of *him* referring to Samuel.

Finally, another potential source of influence on the interpretation of pronouns comes from implicit causality. Some verbs carry with them an implication about what (or who) caused the event. For example, a verb like *amaze* in *The teacher amazed the student* implies causation to the agent—in this case *the teacher*: it was the teacher who did something in the past that then caused the student to be amazed. These can be called NP1 verbs because the causer is the first noun phrase. There are also NP2 verbs like *admire* in which it is the patient (or second noun phrase) who is implied to have done something. For example, in a sentence like *The teacher admired the student*, now it is the student who has done something in the past that then causes the teacher to admire him or her. If the sentence continues with a "because clause" such as "The teacher admired the firefighter because he rescued a trapped man," that clause is read faster when the pronoun can be interpreted as referring to the implicit causer (e.g., fireman) compared with when it cannot.

Garnham, Traxler, Oakhill, and Gernsbacher (1996) looked at how and when implicit causality might influence processing, particularly with pronoun interpretation, by testing two hypotheses: the focusing hypothesis, in which implicit causality would focus attention on whichever referent is the cause of the event encoded by the verb, and the integration hypothesis, in which implicit causality has no immediate effect, but may influence the integration of information about the event. According to the focusing hypothesis, one would expect to find differences in the accessibility of the causer at the verb and then again at the pronoun, but using a probe recognition task, Garnham and his colleagues did not find this. Instead, they found that implicit causality appeared to only have an influence under very specific circumstances, and only at the end of the test sentence. For example, in a sentence like *Walter apologized to Ronald this morning because he had damaged the car*, Garnham et al. did not find any evidence that the verb *apologize* influenced the interpretation of the pronoun *he* toward Walter, until the end of the sentence when all the information from the sentence (in addition to the verb) was available and the end of the sentence caused the need to integration information across the different parts of the sentence. Thus, their results are best accounted for as an effect of integration, rather than an immediate influence of referent accessibility.

In the previous few pages we have explored a number of factors that influence pronoun interpretation (and even production). However, as discussed at the beginning of this section, there are other forms of reference—for example, one can use the phrase "the bird" to refer back to a previously mentioned robin. These types of coreference appear at first glance to be rather different—they provide more semantic information and so the question of identifying the antecedent referent is (apparently) easier. Also, having more information allows for potentially poor matches (e.g., using "the bird" to refer back to a previously mentioned penguin). In fact, there has been a fair amount of research on other forms of coreference, with the general finding that accessibility also influences these forms, and that the typicality of the match between antecedent and coreferring expression is important (under most circumstances, using "the bird" to refer back to a robin is easier that using it to refer back to a penguin). Because of apparent difference in form, there is some debate about whether all types of coreference are subject to the same processing constraints, or whether there is a difference between pronouns and fuller forms of reference. In short, whether pronouns are processed differently than other forms of coreference. Almor (1999) proposed the Informational Load Hypothesis, which presents a unified account of coreferential processing. Other approaches, just as Centering Theory, have focused more on how pronouns are dealt with. Yet another type of approach has proposed that different types of coreferring expressions may be sensitive to different cues or constraints and, further, that this sensitivity may vary cross-linguistically (e.g., Kaiser & Trueswell, 2008).

WRAPPING UP

As we have seen in this chapter, once we start looking at larger contexts and more than one person at a time, things begin to get complicated. However, understanding how language is processed in these cases is vital, because this better reflects how language is used on a daily basis. The results so far suggest that we coordinate closely with other people as we talk and listen to them, but that there are also lots of other factors that still influence how we process language.

REFERENCES

Almor, A. (1999). Noun–phrase anaphora and focus: The informational load hypothesis. *Psychological Review, 106*(4), 748–765.

Ariel, M. (1990). *Accessing noun–phrase antecedents*. London: Routledge.

Arnold, J. E. (1998). *Reference form and discourse patterns*. Unpublished Ph.D. dissertation, Stanford University, Stanford, CA.

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters of choice of referring expression: Everyone counts. *Journal of Memory and Language, 56*, 521–536.

Arnold, J., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language, 51*, 55–70.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal* *of Experimental Psychology: Learning, Memory, and Cognition, 28*(4), 748–769.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language, 42*, 1–22.

Bard, E. G., Lowe, A., & Altmann, G. (1989). The effect of repetition on words in recorded dictations. In J. Tubach & J. Mariani (Eds.), *Proceedings of EUROSPEECH '89* (Vol. 2, pp. 573–576). Paris: European Speech Communication Association.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104*, 437–458.

Branigan, H., Pickering, M., & Cleland, A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*(2), 13–25.

Branigan, H., Pickering, M., McLean, J., & Cleland, A. (2007). Syntactic alignment and participant role in dialogue. *Cognition, 104*(2), 163–197.

Brennan, S. E., & Clark, H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning Memory and Cognition, 22*(6), 1482–1493.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. Li (Ed.), *Subject and topic* (pp. 25–56). New York: Academic Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Clark, H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning & Verbal Behavior, 22*, 245–258.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1–39.

Cleland, A., & Pickering, M. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun–phrase structure. *Journal of Memory and Language, 49*(2), 214–230.

Cornish, F., Garnham, A., Cowles, H. W., Fossard, M., & Andre, V. (2005). Indirect anaphora in English and French: A cross-linguistic study of pronoun resolution. *Journal of Memory & Language, 52*, 363–376.

Cowles, H. W. (2007). The influence of "aboutness" on pronominal coreference. *ZAS Working Papers in Linguistics, 48*, 23–38.

Cowles, H. W., Walenski, M., & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi, 26*, 3–18.

Eberhard, K., Spivey-Knowlton, M., Sedivy, J., & Tanenhaus, M. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research, 24*, 409–436.

Ferreira, V. S., & Dell, G. S. (2000). The effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40*, 296–340.

Fowler, C., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26*(5), 489–504.

Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language, 35*(4), 517–543.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialog—A study in conceptual and semantic coordination. *Cognition, 27*(2), 181–218.

Gernsbacher, M., & Hargreaves, D. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory & Language, 27*, 699–717.

Gernsbacher, M., Hargreaves, D., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory & Language, 28*, 735–755.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation* (pp. 1–68). Cambridge: Cambridge University Press.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition, 13*, 8–20.

Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*(2), 274–307.

Haywood, S., Pickering, M., & Branigan, H. (2005). Do speakers avoid ambiguities during dialogue. *Psychological Science, 16*, 362–366.

Horton, W., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*, 91–117.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.

Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes, 23*(5), 709–748.

Kehler, A. (2002). *Coherence, reference and the theory of grammar*. Stanford, CA: CSLI Publications.

Kertz, L., Kehler, A., & Elman, J. (2006, July 26–29). *Grammatical and coherence based factors in pronoun interpretation*. Paper presented at the Proceedings of the 28th Annual Conference of the Cognitive Science Society, Vancouver, Canada.

Keysar, B., Barr, D., Balin, J., & Brauner, J. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11*(1), 32–38.

Kraljic, T., & Brennan, S. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology, 50*, 194–231.

Malmberg, K. J., & Nelson, T. O. (2003). The word-frequency effect for recognition memory and the elevated-attention hypothesis. *Memory & Cognition, 31*(1), 35–43.

Nicol, J., & Swinney, D. A. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, Special Issue: Sentence Processing, 18*, 5–19.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*, 2382–2393.

Pickering, M. J., & Ferreira, V. (2008). Structural priming: A critical review. *Psychological Bulletin, 134*(3), 427–459.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169–190.

Sedivy, J., Tanenhaus, M. K., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109–147.

Simner, J., & Smyth, R. (1999). Phonological activation in anaphoric lexical access (ALA). *Brain and Language, 68*, 40–45.

Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics 9: Pragmatics* (pp. 315–332). New York: Academic Press.

Stevenson, R. J., Nelson, A. W. R., & Stenning, K. (1995). The role of parallelism in strategies of pronoun comprehension. *Language and Speech, 38*(4), 393–418.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory & Language, 48*, 542–562.

van Gompel, R., & Majid, A. (2004). Antecedent frequency effects during the processing of pronouns. *Cognition, 90*, 255–264.

Wilkes-Gibbs, D., & Clark, H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language, 31*(2), 183–194.

7

Using Your Hands: Sign Languages

Human language isn't limited to spoken (or written) forms, but can also be expressed visually, in sign language. Research on both sign language and how it is processed has been growing quickly over the last decade, with researchers from a number of different fields increasingly interested in it. By studying how sign language differs (and does not differ) from spoken language, researchers are able to better understand how much language is influenced by the medium in which it takes place (e.g., spoken vs. gestural)—this is referred to as a language's "modality." Further, although we now know quite a bit about how spoken and written language is processed, we still don't know how much of how language processing works is due to specific aspects or limitations of the sound medium (or sound medium translated into funny visual squiggles on a page/screen). Sign language thus has a lot to tell us about what parts of language are specific to the type of signal used (e.g., vocal) and what parts are modality independent—that is, which parts of language exist regardless of how you express them. This work also has neuroscientists and evolutionary biologists excited—sign language research can give insights into the neural representations of language and may also give us a larger window into how humans came to have this marvelous ability.

COMMON MISCONCEPTIONS

If you are unsure about what sign language really is, you're not alone. Until William Stokoe's groundbreaking work on the linguistic structure of American Sign Language (ASL) in the 1950s and 1960s (e.g., Stokoe, 1960), even many linguists were unclear about what to think about sign language, or were simply wrong about their characterization of it. Before we can talk about what we know about how sign language is processed, we first need to address two common misconceptions about sign language and understand exactly what sign language is (and is not).

Misconception No. 1: American Sign Language Is Just a Version of English

We could pose this misconception in a number of ways, French Sign Language is just a version of spoken French, British Sign Language (BSL) is just a version of English, and so on. Under this misconception, sign languages are incorrectly seen as just another means of expressing a corresponding spoken language, not unlike the way in which written language corresponds to a spoken language. There are, in fact, manual versions of English in which English words are translated or finger spelled in the exact order you would hear them in English. These are not sign languages but invented sign systems and are referred to as codes—for example, one such code is Manually Coded English.

Sign languages are different in that they are not simply versions of another (spoken) language but are in fact completely independent languages. Sign languages are also independent from each other. In the same way that you would not expect an English speaker to automatically understand someone speaking, for example, Norwegian, you should not expect an ASL signer to automatically understand another sign language. That said, in the same way that spoken languages can share some vocabulary or structural features based on historical associations (e.g., Romance languages have a number of similarities based on their shared language ancestor Latin, as well as ongoing interaction among speakers of these languages), sign languages can also have similarities to other sign languages, based on their shared history. This relationship between spoken and sign languages can be seen when comparing ASL with BSL: despite the fact that English is a dominant spoken language in both the United States and Britain, ASL actually has more in common with French Sign Language than with BSL, due to historical reasons.

Further, in the same way that spoken languages have regional and socially based dialects, so too do sign languages. Saying "y'all" instead of "you" or "you all" or "you guys" tells speakers of English that the speaker is likely from, identifies with, or has spent significant periods of time in the southeastern part of the United States. Such variation can influence what we call things, so that we have *lightning bugs* and *fireflies*, which are in fact the same thing. In the same way, variations in hand shape and other differences can differentiate dialects of sign language.

In short, sign languages are languages in their own right—NOT versions of spoken language.

Misconception No. 2: Sign Language Is Not "Real Language"

This misconception actually comes in a number of flavors, including "Sign language is just pantomime" and "Sign language can't express complex ideas like real language." Of course, this just begs the question—what is "real" language? The answer to this question is more complicated than it sounds and is important to other areas of research, including animal communication and language evolution. It is a fascinating topic in its own right, but a full exploration of it is outside the scope of this book. Instead, we will return to some areas of linguistics that we covered in chapter 2: phonology, syntax, and semantics. Recall that there were two important themes in these areas: separation of form and meaning (e.g., arbitrariness) and hierarchical structuring of form. Importantly, for our purposes here, establishing that sign language is a "real" language means addressing these aspects of spoken language.

ARBITRARINESS

The easiest way to convince yourself that sign language is not simply a series of pantomimes is to watch someone signing and see if you can understand what they're communicating. If you don't know anyone who signs, then there are many native and fluent signers who have video blogs on the Internet that you can watch. It is a telling experience, and I strongly encourage you to do this, preferably now. As with hearing a spoken language that you don't know, it is not only impossible to understand what they are communicating but also difficult in many cases to even tell where one sign ends and the next one begins.

That said, there is a certain amount of iconicity in sign languages—this is the case when a word shares some physical feature of the thing it refers to. An example of this in spoken language comes from onomatopoeia—for example, the words that we use to label the sounds that animals make. What does a dog say? It says woof. A rooster? Cock-a-doodle-doo. And so on. The word for the sound bears a resemblance to the sound itself. A dog doesn't actually say "woof," but English speakers have found a word form that sounds a little like the sound a dog does make when it barks (another sound word!). In spoken language, it would be a little strange to take one of these words and try to make it mean the sound another animal makes. For some reason, we are reluctant to say that a pig meows, even though in principle any set of sounds (e.g., a word) should be able to take on any meaning.

Within a language there are also sound-meaning regularities in which a certain combination (or cluster) of sounds, while meaningless in and of itself, occurs in a set of words that all share a certain similarity. For example, the sound combination "gl" doesn't mean anything in and of itself, and yet there are a number of words in English that start with "gl" that all share a kind of similar idea: gleam, glisten, glitter, glamour, glow, and glimmer. These two phenomena, the matching of sound clusters to meaning and onomatopoeia, are sometimes referred to as *sound symbolism*.

Sound symbolism shows that there are cases in spoken language when sounds are linked in a nonarbitrary way to meaning; however, there are more cases in which there is no particular relationship between the sounds and meaning; indeed, one of the great innovations of language as a communication system is that we can take any sound and apply it to any meaning. This gives us an enormous flexibility in assigning meaning to forms, and lets us build up huge vocabularies out of just a small set of sounds. And so, many language researchers argue, languages must consist of arbitrary mappings between form and meaning. Incidentally, this also means that languages can be mutually unintelligible. If the same set of sounds can mean anything, then different languages are free to assign them to different meanings, and suddenly "ma" can mean mother in one language, and "horse" in another. Not good things to get mixed up. Notice that iconic words, like those for animal sounds, have a certain cross-linguistic intelligibility. What does a rooster say in French? "Cocorico."

This leads us to some of why early characterizations of sign language were so wrong, and why people often have the misconception that sign language isn't actually language. In sign language, there are many signs that do have an iconic aspect to them—they share some feature of the thing that they refer to. However, despite this, many signs still do not have a meaning that is transparent just from watching the sign. Iconicity does not equal transparency. In ASL, for example, the word for *tree* is formed by having the signers arm stand in for a tree—the arm becomes the trunk and the hand and fingers become the bushy top of the tree. However, there is nothing inherent in this sign that requires it to be the word for tree, this same sign could just as easily refer iconically to a spray of water like a fountain or to a road sign—anything with a vertical pole topped by something bigger. Conversely, every object has a number of physical properties that can be incorporated into a sign and so it is still language-specific which features, if any, might be used. Although in spoken language iconic sounds are often intelligible across languages, this is not as true for sign languages. Sticking with the tree example, the sign for tree in Chinese Sign Language involves a hand movement in which the hands could be seen to be wrapped around the trunk of a tree, then slid up along it—it bears no relation in form at all to the ASL sign for tree, yet both are iconic to the extent that they share some features with the thing they label (Klima & Bellugi, 1979). Thus, although both the ASL and Chinese Sign Language signs for tree could be said to be iconic, they highlight different aspects of a tree and so are different from each other. Just as with spoken language, it is necessary to learn the meanings of the forms in order to understand the language—you can't just guess. In this way, sign language may indeed have a higher degree of iconicity than spoken language, but is still much more than a series of pantomime-like gestures. There has been some recent work looking at the influence of iconicity on sign language processing that we will discuss in the second half of this chapter.

STRUCTURE

Let's start with the phonology of sign language. Recall in spoken language that we can break down words into component sounds that have language-specific rules about their formation (phonotactics) and that we can break down the sounds according to a finite set of features that can be combined in different ways (e.g., place of articulation, manner of articulation, voicing). Signs in sign language are built up in the same way—there is a finite set of features or parameters that can be combined to form signs. Further, there are "phonotactic" rules that differ from language to language about how signs may be formed. Instead of parameters coming from points of contact between the articulators in the mouth, in sign language the parameters come from the interaction of hand configuration (or shape), hand orientation, movement, and place of articulation with respect to the body of the signer. There are signs that are distinguished only by a change in one of these parameters. For example, in Israeli Sign Language, the signs for *already* and for *document* differ only in terms of the hand configuration: the movement and position of the sign with respect to the body is identical, but for *already* the hand is held flat with the thumb sticking up while for *document* the hand is closed like a fist (Sandler & Lillo-Martin, 2006).

Next, let's turn to the morphology and syntax of sign language. Just as with spoken language, sign language is more than a series of words strung together at random. There is structure to how words may be combined, and words may be inflected (changed slightly) to reflect the grammatical relationships they share with other words in the sentence. Sign languages take advantage of space in their syntax. For example, recall that in spoken languages there are two basic options for encoding information about who did what to whom. If there is a helping event in which Mary is the helper and Bill is the helpee, in English we use the order of the words to encode this: Mary helped Bill. If we reverse the order, the meaning changes: Bill helped Mary. Other languages add something to the words themselves instead of using word order. In ASL, the information is encoded spatially. For example, in order to encode the same information as before, a signer would establish distinct spatial locations for both Mary and Bill and then would sign the word helped between Mary and Bill—the direction of the sign (from Mary to Bill) would indicate who was the helper and helpee. The use of space in ASL is thus grammatical—this creates an interesting convergence of domains that one doesn't find in spoken languages, the ability to process general spatial information, such as understanding where objects are in relationship to you and each other, and processing *linguistic* spatial information—encoding grammatical information from the location of signs in a grammatical signing space in front of the signer. We will explore this convergence in detail later in this chapter.

Another difference between spoken and sign languages is that many sign languages also use facial expressions to encode grammatical information. For example, in ASL, signers will furrow their eyebrows during the part of the sentence that contains a question. In English, word order is once again used to mark the beginning of a question—*Who did you meet?* versus *I met him*. Here, the *wh*-word that corresponds to who is met (e.g., *him*) is moved to the beginning of the question. So, speakers of English know immediately that they are dealing with a question because they encounter the word *who* at the beginning. Although the word order for these two types of English sentences is quite different, the word order of the corresponding sentences in ASL is the same: the signs for the question are, in order: MEET WHO and the signs for the statement are, in order: MEET HIM. So, based on the signs themselves, comprehenders wouldn't know they were being asked a question until the second sign, WHO *versus* HIM. However, because ASL signers may furrow their eyebrows to indicate the beginning of the question, these two sentences can actually be different right from the beginning: for questions, the signer can furrows his/her eyebrows at the sign MEET. This allows comprehenders to know they are being asked a question from the beginning of the sentence.

As with spoken languages, the rules of how words may be put together and modified differ depending on the language. If these rules are violated, then the sentence is ungrammatical or the word is ill-formed. Also, like spoken language, a full account of sign language syntax is well beyond the scope of this book, but readers interested in knowing more might want to check out books by Emmorey (2002) or Sandler and Lillo-Martin (2006) for more information. Importantly, for our purposes, sign languages have all of the hallmarks of spoken languages—rule-based structures, recursiveness, and so on. With this knowledge, we can now turn to whether the ways in which sign language expression differs from spoken language cause it to be processed differently.

PROCESSING SIGN LANGUAGE

In terms of the rate of speech versus sign, spoken language comes at comprehenders a little faster than sign: speech rates among speakers differ, on average speech contains about 10 to 15 segments per second (Liberman, 1996), whereas sign language contains about 7 to 11 segments per second (Emmorey, 2002). However, both speech and sign have approximately the same rate of information transmitted: both modalities express about 1 coherent unit of information (a proposition) every second (Bellugi & Fischer, 1972). So, in terms of perception, while speech has more low-level units to deal with, the overall amount of information per second is roughly equivalent.

PRODUCTION ERRORS

Another way in which speech and sign are similar during production is that both are prone to errors. Speech errors are mistakes that speakers make when they intend to say one thing but something else comes out instead. Many people are familiar with the idea of speech errors from the notion of Freudian slips, and sitcoms often make use of speech errors for a laugh. Sometimes speech errors do result in something funny or something that listeners feel is telling about the "real" intentions of the speaker. But, in fact, most speech errors are pretty boring. They result from misplacing a unit of speech in some way—for example, a speaker might accidentally swap the initial sounds between two words so that they intend to say "tell a fable" but instead say "fell a table." We know quite a bit about the patterns of speech errors in language and previous work has established that these are not random errors, but reflect something important about how speech is planned and produced (see Levelt, 1993, for more details about both speech errors and language production). For example, we know that speech errors tend to involve similar sounds and happen more often when the speaker is stressed. Researchers have also noticed that there is a "lexical bias" in speech errors such that errors tend to result in real words, like the example given earlier. Errors can occur at any level, so that phonemes, syllables, morphemes, words, and even whole phrases can be swapped, omitted, repeated, and so on. Importantly, speech production errors are *not* errors of knowledge—the errors that you might make while learning a new language because you don't know the right words or rules yet do not count as speech errors (though they may still count as errors as far as your instructor is concerned); also, people who use the wrong word because they are not entirely sure what it means are not making speech errors. If you *intend* to say that your shirt is inflammable to mean that it cannot be set on fire, this is not a speech error.

So, what about sign language? Actually, production errors happen in sign language production in much a similar way to spoken language. Newkirk, Klima, Pederson, and Bellugi (1980) reported errors in signing that involve exchanges of different parameters, such as hand shape, place of articulation, and movement. For example, a signer might intend to sign MUST SEE, whose signs involve different hand shapes, and during production might accidentally use the hand shape of SEE in the sign for MUST, a "speech" error. As with spoken language errors, sign language errors almost always result in words that are possible in the language. However, there are some differences in the patterns seen in sign language production errors, which are likely due to modality differences. For one thing, evidence from German Sign Language suggests that signers are likely to notice and repair their errors much sooner than speakers (Hohenberger, Happ, & Leuninger, 2002), which may be due to the slightly slower production rate of signers. Another difference is that the lexical bias effect seen in speech has not been found in sign language (Emmorey, 2002); Emmorey points out that despite its robust effect in English, the lexical bias effect isn't always seen even in other spoken languages: in one study, del Viso, Igoa, and Garcia-Albea (1991) did not find a lexical bias effect in Spanish. They suggest that perhaps the rate of speech had some effect here (Spanish tends to be produced a bit faster than English), but sign language's relatively slower rate would argue against this possibility.

So, in terms of production errors, sign language and spoken language are remarkably similar yet do have systematic differences. This suggests that while the system underlying the production of sign and speech is probably very similar, modality-dependent differences do exist.

CATEGORICAL PERCEPTION— THE NEED FOR SPEED REVISITED

As we have seen several times in this book, speech is very rapid and any comprehension system needs to be able to handle that rapidly changing acoustic information. One way that the system handles this is via categorical perception. As the name suggests, this means that we perceive something categorically. But what? And why does it matter? There is a lot of evidence to suggest that hearers perceive consonant sounds in language categorically. In practical terms, this means that when people hear a consonant, it is rapidly assigned to a phonemic category, which depends on one's language. Importantly, this ability to rapidly (and correctly!) identify a consonant means that hearers loose the ability to distinguish between sounds that fall within the same category. In fact, this is the very definition of categorical perception used by some textbooks: categorical perception is the inability to perceive the difference between two sounds that fall within the same phonemic category. The classic example of this is how hearers distinguish voiced and voiceless stop consonants—how they tell the difference between /b/ and /p/ or between /k/ and /g/. Crucially, stop consonants don't have any sound of their own—by definition they are produced by stopping airflow through the vocal tract. However, their identity can be determined by what comes immediately before and after—the transitions of the consonant. Hearers use something called *voice onset time* (VOT) to determine voicing. If the time between when the airflow is stopped to when the vocal chords begin to vibrate is shorter than about 35 ms, then hearers perceive the sound as voiced—for example, /b/. If that space of time is longer, then the consonant is perceived as voiceless—/p/. That space of time is the VOT. There are few things to note about this process. First, there is no in-between sound that people report hearing—it is either a /b/ or a /p/, not an "almost" /b/ or "kind of" a /p/. Second, there is very little time where people are uncertain about the category—it is not the case that when there is a VOT of, say 45, that people report hearing a /p/ more than when the VOT is 65. Finally, and here is the most important part, if you play two sounds back to back and ask participants to say whether the sounds are the same or different, their ability to distinguish between the two sounds depends on whether the difference between the two crosses the category boundary. So, if you were to play two sounds, one with a VOT of 10 and one with a VOT of 30, participants would not report hearing those as different sounds—they would say that they are same. The same is true if you play sounds with VOTs of 40 and 60. But, if you were to play a sound with a VOT of 20 and one with a VOT of 40, then people are much more likely to report that they hear these as two different sounds. The spacing between the sounds along the VOT parameter is the same as in the first two cases, but now people can hear the difference. Why is this? It's because the two sounds cross the phonemic barrier.

So, what about sign languages? Do they show a similar processing effect? This is not a test of language-hood—we know sign language is a real language. Rather, it is a test of whether there is something particular about speech that requires (or at least has produced) categorical perception in order to comprehend it. One reason to think categorical perception is not necessary in sign language is the relative difference in speed between the two modalities—recall that the rate of transmission for sign language is slightly slower than for speech. Even within speech, not all sounds are categorically perceived; vowels, which last longer on average than consonants, don't appear to be perceived categorically. One theory is that the rapid nature of consonants requires categorical perception. Another reason that sign may not require (or at least produce) categorical perception is that this process may be limited to sound perception rather than visual perception.

Most of the research in this area has focused on the categorical perception of hand shape (e.g., Baker, Idsardi, Golinkoff, & Petitto, 2005; Emmorey, McCullough, & Brentari, 2003). In one study, Emmorey et al. (2003) investigated categorical perception in ASL in experiments that compared both native Deaf signers of ASL with hearing nonsigners that had no prior experience with ASL. In the first experiment, the researchers used an ABX discrimination task and a categorization task. In the ABX discrimination task, participants are presented with three stimuli, one at a time. Participants must decide whether the last stimulus (X) is the same as the first (A) or second (B) stimulus.

The A and B stimuli are either from the same phonemic category (e.g., both types of /b/) or different categories (e.g., /b/ and /p/). In the categorization task, participants are presented with two category end points and asked to decide whether a particular stimulus belongs to one category or the other. In both tasks accuracy is analyzed. In this study, instead of looking at speech sounds and manipulating something like VOT, the researchers presented images of signs and manipulated hand shape and place of articulation. For example, one hand shape in ASL is with the fingers straight and held together so that the hand is perfectly flat. A contrasting hand shape has the same configuration except that the fingers are curled into something like a fist. Between the completely flat finger position and the completely curled position there are a lot of possible partially curled positions; the question is how are these in-between positions perceived? If perception of these two hand shapes is categorical, then there should be some state of partial finger curl in which perception of the hand shape switches from one sign to another, in the same way that a voiced stop is suddenly heard as a voiceless stop when the VOT increases past 35 ms. Emmorey et al. (2003) found that in the categorization task both deaf native signers and hearing nonsigners had similar responses—both showed a sharp boundary between the two end point hand shapes (fully flat vs. fully curled) in which categorization changed from one category to the other. However, in the ABX discrimination task (as in, can you tell these two hand shapes apart?), only the native signers showed the classic categorical perception pattern: signers had the highest accuracy in this task when the A and B images were from different hand shape categories. This was not true for the nonsigners. Interestingly, both signers and nonsigners showed a category boundary effect in this task for stimuli that differed by place of articulation. Any effect that is shared by both groups cannot be accounted for strictly on linguistic grounds because the nonsigning participants have no linguistic knowledge of ASL. Thus, because the signers alone show evidence of increased perception accuracy when two stimuli cross a hand shape category boundary, these results suggest that sign language comprehenders do show categorical perception that is based on language experience, at least for hand shape.

LEXICAL ACCESS IN SIGN

In chapter 4, we talked about some of the factors known to influence how quickly people are able to retrieve words from their mental dictionaries—termed lexical access. Four factors were specifically mentioned: frequency, age of acquisition, neighborhood size, and semantic priming. Researchers have also investigated the effect of these factors in sign language and have found similar results to those found in spoken language. The frequencies of signs for a given sign language are not readily available the way that they are for many spoken languages; however, some work has looked at a related factor, word familiarity. Unlike frequency information, which is on word counts from large sets of written and spoken language corpora, familiarity information is based on rating data from speakers about how familiar or unfamiliar a particular word is. Nonetheless, despite these differences between the two measures, both high-frequency and highly familiar words are responded to faster than words that are lower in frequency and/or less familiar. Signs that are more familiar are also responded to faster than less familiar signs (e.g., Carreiras, Gutierrez-Sigut, Baquero, & Corina, 2008). Age of acquisition for particular words has not been directly studied, but one study on the influence of the initial age of acquisition for ASL more generally found that reaction times for sign recognition increased with age of acquisition (Mayberry & Witcher, 2005). Neighborhood size also has a similar effect in both modalities. Recall that the more neighbors (similar sounding words) a word has, the more quickly it is responded to in spoken language. Carreiras et al. (2008) looked at neighborhoods based on hand shape in Spanish Sign Language and found the same effect—more neighbors meant faster and more accurate responses. Semantic priming has also been found, in one study looking at ASL, Corina and Emmorey (1993) found faster response times for sign pairs that were semantically related.

In chapter 4, we also discussed in some detail the debate about how ambiguous words are handled. Little work has been done in this area in sign language, perhaps in large part because sign languages appear to have relatively low levels of ambiguity. Emmorey (2002) reports a comparison of rates of ambiguous words in English, Navajo, ASL, and BSL and finds that about 20% of English words have multiple meanings, but only about 5% of all three other languages do (4%, 6%, and 7%, respectively). The fact that Navajo, a spoken language, also has a relatively low rate of ambiguity suggests that the amount of lexical ambiguity in a language may not be dependent on modality—spoken languages may have greater or fewer numbers of ambiguous words. Further work in this area is needed before we can know whether modality does play a role in ambiguity rates.

BENEFITS FOR ICONICITY IN LEXICAL ACCESS

Returning to the issue of iconicity in sign languages, some recent work has examined whether the language processing system in native and fluent signers can take advantage of the more transparent link between form and meaning in many signs (Thompson, Vinson, & Vigliocco, 2009). Thompson and colleagues looked specifically at signs that share features with the objects they refer to, and in particular they were interested in whether native signers would be faster to recognize that a sign referring to an object when the shared features between object and sign were highlighted. The researchers also looked at both proficient second language learners of ASL and native speakers of English who had no previous experience with ASL.

In the critical trials of this task, participants would see a picture of an object followed by a short video of someone producing the sign for that object. There were two pictures of each object tested, each highlighting a different aspect of the object. For example, one critical object was a hearing aid: both pictures showed a profile of a person wearing a hearing aid, but in one picture the hearing aid was one that loops over the back of the ear, in the other the hearing aid was an ear bud type that doesn't have a loop. The sign in ASL for hearing aid includes a hand shape that shares the loop feature of the hearing aid. Thus, would native signing participants be faster to confirm that the picture of the hearing aid and sign for hearing aid were referring to the same thing when the picture of the hearing aid had the loop than when it didn't? The answer to this is yes—in fact, both native signers and signers who learned ASL as a second language were faster. However, this could be due to some other reason—perhaps people just associate particular pictures with their object more easily and it is not iconicity that is behind the results. So, it is important to know what the nonsigning participants did. For their trials, instead of playing a video of someone signing the word for the object, the researchers played a video of someone saying the word in English. For these participants, there was no difference in response times between the pictures. These results suggest that iconicity in signs may actually be an advantage for lexical access in sign language—shared features between an object and its sign may allow faster access to those object features.

LANGUAGE IN SPACE

There are two topics that we'll cover in this last part of the chapter, both having to do with the spatial nature of sign languages. Recall from earlier in the chapter that orientation in space, rather than word order, determined the meanings of sentences and phrases in ASL. The lack of reliance on word order is not particular to ASL or sign language—there are many spoken languages that also use some other means to encode things like who did what to whom. However, other spoken languages don't use space to do it. So, in this section we'll explore the use of space in sign language in a little more depth.

GRAMMATICAL SPACE VERSUS SPATIAL REPRESENTATIONS

Close your eyes for a moment and picture a room from someplace that you've lived recently. If asked, you could describe the room in some detail: The chair is to the left of the couch, both facing toward the center of the room; there is rug underneath them. And so on. This ability is based not just on memory, but on your ability to orient yourself and other objects in space—to construct spatial representations of the world around you. This is a separate ability from your linguistic ability to describe the spatial representation itself using speech. So, what about sign language? Because sign language crucially uses spatial information to convey linguistic information, should it be the case that in sign language users these two abilities *are* related? As it turns out, the answer is no. While studies of the visuospatial skills of signers have shown increases in specific skills, such as memory for spatial locations (see Emmorey, 2002, for an overview), the use of space in sign language remains independent from nonlinguistic visuospatial skills. Poizner, Klima, and Bellugi (1987) distinguish between spatial mapping and spatialized syntax in sign language. The first is the use of space in describing visual scenes and the other is the use of space for purely linguistic functions, such as for coreference or verbs that require particular spatial configurations. Several studies looking at patients with brain trauma have found dissociations between these skills, such that patients may have intact language ability and sign using grammatical space perfectly, yet what they describe using that sign syntax is incorrect. Inversely, patients may have lost language ability, yet still be able to demonstrate knowledge of spatial relationships.

COREFERENCE

In the previous chapter, we discussed coreference in spoken languages and found that different forms of reference have differing impacts on the mental representations of a discourse and processing sentences in that discourse. ASL also makes a distinction between descriptive nouns such as *the principal* and pronouns to refer to the same thing, such as *he*. ASL has a further distinction that English doesn't have, although other languages like Spanish and Chinese do—the distinction between explicitly mentioned (overt) pronouns and unmentioned (null) pronouns. In ASL, overt pronouns are formed by pointing to the space that represents the thing being referred to. Null pronouns occur with certain verbs in which the sign for the verb itself is made in the space of the referent. For example, if one wants to say *Bill blamed Jill*, they would sign BLAME starting in the syntactic location of Bill and moving to the syntactic location of Jill. This would be in some ways equivalent to *He blamed her*, but without any overt pronoun (he or she) explicitly mentioned.

Recall that one of the questions addressed in research on coreference is whether the representation of an antecedent is activated by encountering a pronoun, and that the answer appears to be that it does. Research on coreference in ASL has shown that this is also true for both overt and null pronouns in this language (Emmorey & Lillo-Martin, 1995), which is particularly interesting given the difference between these two types of pronouns—in the null case, simply signing the verb in the space of the antecedent was enough to cause faster response times to a probe word corresponding to that antecedent.

WRAPPING UP

In this chapter, we have seen how research on sign languages can show which aspects of language are due to language itself and which are due to the modality in which the language takes place. As with the study of any language, findings from research on sign languages will help us gain a better understanding of how we are able to communicate complex ideas so effortlessly. Further, studying the similarities and differences between any two languages helps us better understand what the universal underpinnings of language processing are, and what are determined by the specific features of a particular language. This benefit has the potential to be even larger when comparing sign and spoken languages because of the large difference of modality between them. And so, current research in sign language continues to investigate the similarities and differences between the forms of signed and spoken languages as well as the factors that influence their processing. Many researchers who are interested in the relationship between language and other cognitive processes, as well as language and the brain, have also begun to focus on sign language. The current pattern of results suggests that although spoken and sign languages are processed in remarkably similar ways, there are important differences that appear to be due to their modality.

REFERENCES

Baker, S., Idsardi, W., Golinkoff, R. M., & Petitto, L. (2005). The perception of handshapes in American Sign Language. *Memory & Cognition*, *33*(5), 887–904.

Bellugi, U., & Fischer, S. (1972). A comparison of sign language and spoken language. *Cognition*, *1*, 173–200.

Carreiras, M., Gutierrez-Sigut, E., Baquero, S., & Corina, D. (2008). Lexical processing in Spanish Sign Language (LSE). *Journal of Memory & Language, 58*(1), 100–122.

Corina, D., & Emmorey, K. (1993). *Lexical priming in American Sign Language*. Poster presented at the 34th annual meeting of the Psychonomics Society, Washington, DC.

del Viso, S., Igoa, J. M., & Garcia-Albea, J. E. (1991). On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research, 20*(3), 161–185.

Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Emmorey, K., & Lillo-Martin, D. (1995). Processing spatial anaphora: Referent reactivation with overt and null pronouns in American Sign Language. *Language and Cognitive Processes, 10*(6), 631–664.

Emmorey, K., McCullough, S., & Brentari, D. (2003). Categorical perception in American Sign Language. *Language and Cognitive Processes, 18*(1), 21–45.

Hohenberger, A., Happ, D., & Leuninger, H. (2002). Modality-dependent aspects of sign language production. Evidence from slips of the hands and their repairs in German sign language (Deutsche Gebärdensprache [DGS]). In K. Cormier, R. Meier, & D. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken language* (pp. 112–142). Cambridge, MA: Cambridge University Press.

Klima, E., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Levelt, W. J. M. (1993). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Liberman, A. (1996). *Speech: A special code*. Cambridge, MA: MIT Press.

Mayberry, R., & Witcher, P. (2005). *What age of acquisition effects reveal about the nature of phonological processing*. San Diego, CA: University of California.

Newkirk, D., Klima, E., Pederson, C., & Bellugi, U. (1980). Linguistic evidence from slips of the hand. In V. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen and hand* (pp. 165–198). New York: Academic Press.

Poizner, H., Klima, E., & Bellugi, U. (1987). *What the hands reveal about the brain*. Cambridge, MA: Bradford Books/MIT Press.

Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge, UK: Cambridge University Press.
Stokoe, W. (1960). *Sign language structure: An outline of the visual communication systems of the American Deaf, Studies in Linguistics: Occasional Papers* (Vol. 8). Buffalo, NY: University of Buffalo.
Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2009). The link between form and meaning in American sign language: Lexical processing effects. *Journal of Experimental Psychology: Learning Memory and Cognition, 35*(2), 550–557.

8

How Good Is "Good Enough"?

Much of what we have talked about in this book has assumed that comprehenders are trying to process the language that they hear or read as fully as possible—that they interpret each word completely and build complete syntactic structures (sometimes even more than one) when they encounter sentences. But, recent work has cast some doubt on just how complete these representations really are.

Suggestive evidence comes from the early 1980s and the Moses illusion (Erickson & Matteson, 1981), named after one of the most commonly cited sentences that exemplifies it. This illusion centers on how people initially respond when asked to say whether certain sentences are true or false, such as:

1. Moses put two of each sort of animal on the Ark.

Most people, when they first encounter this sentence, answer that it is true. However, the problem is that it was not Moses, but rather Noah who put animals on the Ark. So, how do people who are fully familiar with the story of Noah and the Ark fail to notice this? Clearly, they are not paying attention to all the details of the question. Other similar effects are found when people are asked questions like, "After an air crash, where should the survivors be buried?" Barton and Sanford (1993) found that half of participants who were asked this question responded with an answer that indicated that they should be buried where their relatives wished them to be buried. These participants failed to notice that survivors are living people.

Another example of something less than full processing comes from the typical interpretations of sentences like "No head injury is too trivial to be ignored" (e.g., Wason & Reich, 1979). This is usually interpreted to mean that even if a head injury seems very trivial, it should still be treated. However, if we compare it with another, similar sentence, we can see the problem: "No knife is too small to be banned." Here, this seems to mean the same thing: even if a knife seems very small, it should still be banned. But, here we now see the problem: in the first case, the interpretation means the opposite of what the sentence actually says (using *treated* as opposed to *ignored*). Even if a head injury seems very trivial, it should still be *ignored*. The sentence actually doesn't make very much sense, from a real-world perspective. But, we ignore the particular grammatical details and interpret it to mean something sensible.

All of these cases involve some pretty complicated situations. For the Moses illusion and the survivor case, the problem hinges on not paying attention to the meaning of a particular word with respect to the bigger context. But, they also depend on trusting the assumptions of the question or statement, particularly in the survivor question: the question presupposes that survivors should be buried, and so people may sometimes be lulled into going along with this. However, the larger point that the responses to these types of sentences make is intriguing: if there are times when people are not fully processing language, when are these times? Essentially, when else are people not fully processing language? This is the question that this chapter deals with.

This is not a trivial issue—until recently it was taken as established and assumed by most theories of language processing that people process language as fully as possible, in real time. The debates in language processing have really centered on what kinds of information or cues get used during processing (and when they get used), not whether people are not actually processing language fully. Thus, if it really is the case that people are not always processing language fully, but rather relying on shallow processing strategies to create "good enough" representations of what they are reading or hearing, then we will need to reconsider almost every theory of how language is processed in real time.

Christianson, Hollingworth, Halliwell, and Ferreira (2001) addressed this by questioning the assumption that after recovering from a garden path sentence, readers fully reinterpret and thus completely repair their initial, incorrect syntactic representation to reflect the new, correct interpretation. In a series of experiments, they presented participants with sentences like the following one, among others:

2. GP, Plausible: While Bill hunted the deer ran into the woods.
3. GP, Implausible: While Bill hunted the deer paced in the zoo.
4. Non GP, Plausible: While Bill hunted the pheasant the deer ran into the woods.

Crucially, in the plausible case, the initial misinterpretation, for example, that Bill hunted the deer is compatible with the deer running into the woods, whereas in the implausible case it is not compatible with the deer pacing in the zoo. The researchers also had longer versions of each of these sentences, in which more words were added to postpone the disambiguating word. (While Bill hunted the deer that was brown and graceful ran into the woods.)

They had participants read a sentence and then answer a yes/no question such as "Did Bill hunt deer?" They were then asked to rate their confidence in their answer. For critical trials, the question targeted the initial misinterpretation. The researchers were specifically interested in whether they would get more incorrect responses depending on the type of sentence. They found this was indeed the case: participants were less accurate in the garden path sentences, and further, were less accurate when the ambiguous region was longer and when the misinterpretation was plausible. In the plausible, long condition participants responded roughly at chance—giving an incorrect answer 51.2% of the time. The short plausible was not much better, with 44% incorrect. The most accurate conditions were the non-garden path short and implausible short, with 21.4% and 20.2% errors, respectively. But, how confident were people in their answers? It may well be the case that people are scoring poorly but know that they may well be wrong in many cases. However, across all conditions, participants were pretty confident in their answers. This pattern of results was found regardless of whether the participants read the whole sentence at once or word by word.

One objection is that perhaps participants in this experiment didn't actually reinterpret the garden path sentence—that is, perhaps they have an incorrect interpretation so often because they didn't actually ever get the right interpretation. That would be an interesting outcome in and of itself, but in a follow-up experiment in which the researchers also asked about the second part of the sentences (e.g., Did the deer run into the woods?), they found that participants were very accurate on these questions while simultaneously very inaccurate about the first questions. This suggests that their participants were actually holding two syntactically contradictory interpretations in their heads at once. That is to say—Bill hunted the deer AND the deer ran into the woods. Thus, while the participants do correctly interpret "the deer ran into the woods," they still retain the original misinterpretation that Bill hunted the deer. Interestingly, how likely they were to keep this misinterpretation depended on how plausible the two events combined were (implausible events prompt better accuracy) and how long the misinterpretation was possible (with shorter lengths prompting better accuracy). However, even "better accuracy" means error rates of 40% and 50% across different experiments!

Another objection, however, is that verbs like *hunt* may encourage interpreting the following noun as the object, even if it is not part of the information that is explicitly provided syntactically. The researchers thus ran two further experiments using a special class of verbs that are well known to require a reflexive interpretation in that there is no object given. What does this mean? These verbs (called reflexive absolute transitive, or RAT, verbs) generally deal with actions related to personal hygiene, such as wash, bathe, shave, and dress. If someone says "Mary bathed" then we interpret the unsaid object as a reflexive (*herself*). So, Mary bathed means Mary bathed herself. If there is an object given, then the subject does the action on that verb: Mary bathed the baby means that Mary did not bath herself, but gave a bath to a baby. These verbs are useful here because if there is no object provided in the sentence, people should give the reflexive interpretation. So, in "While Anna dressed the baby that was small and cute spit up on the bed," Anna dressed herself and the baby spat up. They compared these types of verbs with verbs like before, in which the object is optional, but does not have a reflexive interpretation. However, even for these RAT verbs, participants gave incorrect "yes" answers to questions like "Did Anna dress the baby" 65.6% of the time. The optionally transitive verbs were even worse—with 75% incorrect answers.

These are intriguing results—could it be that people were really not completely correcting their initial mistakes while reading? It appears so. Follow-up work showed that this effect was more robust for older adults than younger adults (Christianson, Williams, Zacks, & Ferreira, 2006). Of course, this opens up a whole new issue—are people completely interpreting what they read the first time?

The sentences in Christianson et al. (2001) are, in some cases, pretty complicated, and all of the critical sentences were ambiguous. Perhaps this effect, as striking as it is, is limited to complex or ambiguous sentences. It is plausible that people may process less fully when in a challenging linguistic environment. This possibility was addressed by Ferreira (2003). She tested people's ability to assign roles to the referents in very simple sentences. For example, if we have a sentence like "The mouse ate the cheese" we know that the mouse did the eating and it was the cheese that was eaten. In linguistic terms, the verb *eat* assigns two roles—an agent role (to the subject, mouse) and a theme or patient role (to the object, cheese). Ferreira tested for whether people were able to accurately assign roles in these simple sentences by having the participants listen to the sentences and then respond verbally to a prompt word that asked to them to say who/what was the "acted-on" thing in the sentence or what was the "do-er" in the sentence. So, if you heard the sentence above and then saw "Do-er" the correct answer would be for you to say "mouse." Ferreira coded accuracy and also looked at how long it took people to respond. She was particularly interested in whether people would be less accurate for relatively simple, unambiguous sentences that nonetheless had the order of the agent/patient reversed. In the first experiment, she compared active and passive sentences. The passive version of our earlier sentence would be "The cheese was eaten by the mouse." She also looked at sentences like these that were nonreversible (e.g., cheese cannot eat mice), sentences that were reversible but implausible if reversed, and sentences in which either referent could be the agent or patient/theme. She gave these sentences in both active and passive versions, and both active and passive tested assigning the nouns to both roles. Examples of these sentences in active and passive are as follows:

5. Nonreversible, Plausible: The mouse ate the cheese/The cheese was eaten by the mouse.
6. Nonreversible, Implausible: The cheese ate the mouse/The mouse was eaten by the cheese.
7. Reversible, Plausible: The dog bit the man/The man was bitten by the dog.
8. Reversible, Implausible: The man bit the dog/The dog was bitten by the man.
9. Symmetrical, Version 1: The woman visited the man/The man was visited by the woman.
10. Symmetrical, Version 2: The man visited the woman/The woman was visited by the man.

Ferreira found that while overall accuracy was high, as one might expect, there were still some interesting patterns in the data, as well as some conditions in particular that were surprisingly low. For all three sentence types, passives took longer to answer and were answered incorrectly more than actives. For example, in reversible, biased sentences, in which semantic information can help decide who did what to whom, answers to passive were correct only 81% of the time (compared with 99% for actives). In the implausible condition, when the agent of the passive was probed, participants responded accurately only 72% of the time. Also, people were more accurate at questions that probed the first, subject noun in either condition.

From a certain perspective, passives are more complicated than active sentences and so perhaps it is the case that passives are more difficult simply because they are more complicated. Ferreira addressed this by comparing passive with subject clefts like the following ones:

11. It was the mouse that ate the cheese.
12. It was the dog that bit the man.

13. It was the man that visited the woman.

This type of construction, while unambiguous, was also somewhat more complicated than a simple active sentence, yet when it was used in place of active sentences in a second experiment, passives were still responded too more slowly and less accurately. It appears that the important difference between subject cleft and actives on one hand, and passives on the other, is that the order of the roles is reversed between them: in active sentences, the agent comes first. This is also true for subject clefts. But, in passives, the agent comes last (if it is mentioned at all). This predicts that it is not just passives that should cause poorer comprehension, but any construction that reverses the order of the roles. A final experiment in which subject clefts and object clefts are compared confirms this: subject clefts and actives are both easy and passives and object clefts are difficult. Ferreira argues that comprehenders do not necessary fully process language as they encounter it, but instead rely on heuristics to provide a "good enough" representation. In English, it is so often the case that the noun before the verb is the agent of the verb and that any noun following the verb is *not* the agent, that English speakers can simply assign agent status to the first noun. Of course, accuracy in her study for passives was not 0%, and so clearly people are able to assign the correct role to the correct noun, but Ferreira argues that this is a secondary effort—if the comprehender notices that their initial "good enough" interpretation is, in fact, not good enough because it is incorrect, they go back and reanalyze the sentence more fully.

This, in turn, suggests that the degree to which people bother to process language fully could be under strategic control. This does not mean that we consciously decide to process or not process fully (though it could). Instead, perhaps we only process language as fully as necessary for the needs of the current communicative situation (Ferreira, Bailey, & Ferraro, 2002). There are number of studies that support this idea, showing for example that people may not select a particular sense of a word if it is not important (e.g., Frazier & Rayner, 1990; Frisson & Pickering, 1999) and that certain types of pronouns may not be fully interpreted if they do not need to be (Koh, Sanford, Clifton, & Dawydiak, 2008; Poesio, Sturt, Artstein, & Filik, 2006). Using a change-detection task (Sanford & Sturt, 2002), Koh et al. (2008) found that readers more frequently failed to notice a difference between two otherwise identical texts when a pronoun was changed when that change did not result in a contradictory meaning. For example, if a two people go to a mall and they find a parking space, then it follows that each individual found a space. Readers were less likely to notice a change when "they" found a parking space in the first version of the text but "he" found a parking space in the second. These results suggest that readers fail to distinguish (or underspecify) between whether an action was carried out by one individual on behalf of the group, or whether each member of the group performed the action.

Further evidence in support of good enough hypothesis comes from Swets, Desmet, Clifton, and Ferreira (2008), who tested specifically whether task demands could influence depth of processing. They focused on a previous finding concerning the attachment of ambiguous clauses, such as in (14) to (16):

14. The maid of the princess who scratched herself in public was terribly humiliated.
15. The son of the princess who scratched himself in public was terribly humiliated.
16. The son of the princess who scratched herself in public was terribly humiliated.

The first sentence is fully ambiguous with respect to who did the scratching: it could be the maid or the princess. The other two sentences are not ambiguous, but show that the reflective *himself* or *herself* could be interpreted as corresponding to either the first or second noun (e.g., maid/son or princess). Previous work (Traxler, Pickering, & Clifton, 1998; van Gompel, Pickering, & Traxler, 2001; van Gompel, Pickering, Pearson, & Liversedge, 2005) found that with sentences like these, readers actually took a shorter amount of time to read the ambiguous sentence compared with the unambiguous versions. The details of the various accounts of this effect are not important for our present purposes, but they assume, in line with accounts of other phenomena, that readers attach the ambiguous phrase to rest of the sentence when they encounter it. But, perhaps readers don't actually do this—perhaps they do not try to interpret who scratched themselves unless they have to. Swets et al. (2008) wanted to test this hypothesis: that task demands (e.g., knowing that attachment was required or not required to correctly answer questions about the sentences) would influence whether readers actually did bother to interpret (or attach) phrases like "who scratched herself" to the *maid* or *princess*. That is, would readers underspecify the representation of the sentence if there was no reason to fully interpret it? They had people read the same sets of sentences, but for some readers all of the questions required a full interpretation of the sentence (e.g., Did the princess scratch in public?) while for other readers the questions were more superficial (e.g., Was anyone humiliated?). Swets et al. (2008) found that the pattern of reading did in fact change depending on the type of questions asked. Not only did the more specific questions prompt more careful reading, but it appeared to shift strategies for ambiguity resolution: when asked the full interpretation questions, readers spent more time reading the region immediately following the reflexive when it referred unambiguously to the subject (maid/son) than when it was either ambiguous or referred unambiguously to the second (N2) noun. For the superficial questions, they found the same pattern as earlier studies—with fastest times for the ambiguous sentence and no difference between N1 and N2 attachment. These results suggest that readers may only interpret sentences as fully as necessary for the task, and supports the idea that we may have underspecified or "good enough" representations of language unless it is necessary to create a more detailed interpretation of a sentence.

Why have "shallow" or underspecified processing? Sanford (2002) suggested that a system with finite resources should be able to allocate those resources flexibly—that is, just as in other systems (e.g., like vision) we can't pay attention to everything all the time, and so we process fully only what we need to. Indeed, there is a growing body of evidence that languages allow speakers to structure their utterances in a way that can flag certain parts of the sentence as particularly important or worthy of special attention. The particular form of an utterance, independent of its meaning, can provide cues about what part of the sentence contains previously known or inferrable information and what part of the sentence contains new or uninferrable information. This is the information structure of a sentence and linguists have long been interested in both how languages encode information structure and how information structure interacts with meaning (Lambrecht, 1994). Recently, psycholinguists have been interested, too, in how information structure influences language processing.

For our present purposes, a key question is how language might direct attention to particular information that is of interest. If people really are generally processing language in a shallow way unless otherwise needed, then it would be very good to know what types of circumstances cause "deeper" processing. Swets et al. show that task is an influence—knowing that you need to know a particular piece of information. But, what about the sentence or text itself?

Several studies have looked at the influence of particular syntactic forms on focusing attention. For example, Birch and Garnsey (1995) used two types of constructions to focus attention, *it*-clefts and *there*-insertions, which are shown in (17) to (19) with the key position underlined. The standard, active version of the sentence is given for comparison.

17. Standard active: The boy kissed the girl.
18. It-cleft: It was <u>the boy</u> who kissed the girl.
19. There-insertion: There was <u>a boy</u> who kissed a girl.

They conducted a questionnaire study that confirmed that readers consider the target word to be the most important word in the sentence more often when it was in one of these focusing positions compared with when it was not. They also found that participants were both faster and more accurate at recognizing a word (e.g., *boy*) when it had been in these types of constructions in a previous sentence compared with when it had not. Further, they also found that these focus positions appeared to make phonological information about the focused element more readily available, but did not find clear results for semantically related information. From this, the authors suggest that it-cleft and there-insertion structures emphasize referents, but they may only highlight particular aspects of them, perhaps those that are specific to the context they are used in.

Sturt, Sanford, Stewart, and Dawydiak (2004) used a change-detection paradigm in which people would have to report whether there was a difference between two sentences. Participants would first read a short passage and then, after a half-second pause with a blank screen, see a passage that was either identical or differed by a single word. Participants were asked to say whether they had noticed a change, and if so, what it was. In critical passages, the changed word was either in a clefted syntactic position or it was not. The changed word could also either be closely related to the original word (e.g., *beer* for *cider*) or unrelated (e.g., *music* for *cider*). An example of a passive with the changed word in a clefted position is given below, with the changed word underlined.

20. Everyone had a good time at the pub. A group of friends had met up there for a stag night. What Jamie really liked was the <u>cider</u>, apparently.

Sturt et al. found that when the changed word was semantically unrelated, the position of the word in the sentence made no difference; participants were very good at noticing changed words when they differed clearly in meaning. However, when the changed word was semantically related to the original word, then participants were significantly more accurate at detecting the change when the word was in a clefted position. Sturt et al. found the same effect when they manipulated focus using prior context, comparing (21) with (22) as context sentence for the target sentence, (23), in which the changed word is underlined.

21. Everybody was wondering which man got into trouble.
22. Everybody was wondering what was going on that night.
23. In fact, the man with the <u>hat</u> was arrested.

These results also suggest that focusing constructions like clefts, as well as focusing contexts in which d-linked *wh*-questions are used, cause more attention to be paid to the focus. In fact, this effect is not limited to clefts or discourse structure—even just placing the key word in italics can cause increases in change detection (Sanford, Sanford, Molle, & Emmott, 2006). Sanford et al. (2006) also found that placing emphasis on the target word in auditory materials would also increase change detection. This type of focusing of information can even make the Moses Illusion less pronounced. In looking at the effect of focus status on the effectiveness of this illusion, Bredart and Modolo (1988) found that placing the incorrect information (e.g., Moses, who did not put animals in an ark) in a cleft caused people to catch the error more often.

WRAPPING UP

In this chapter we've seen that we process language less fully than previously thought, but that this processing depth appears to be dynamic—we can process language more completely when the task requires it, or if our attention to drawn to a particular part of the sentence by linguistic, pragmatic, or even typographical cues. From previous chapters we have seen that there is plenty of evidence showing that we use all kinds of cues to construct structure and meaning from the input that we receive, but it may well be the case that we only fully engage in this process under certain circumstances. Otherwise, we may construct representations that are underspecified until they need further elaboration.

WHERE DOES THIS LEAVE US?

In this book, we've explored a number of key topics in psycholinguistics and covered many experimental findings. Where do all of these details leave us in terms of the bigger picture of how we (seemingly) effortlessly process something as complex as language? Work on ambiguities in language has shown us that information about how plausible something is and how frequently we encounter words helps us avoid misunderstandings and recover faster when they do occur. What is still not entirely clear is whether we can avoid being led down the garden path altogether, although the current body of evidence suggests that we can, at least some of the time. Context appears to play an important role in language processing both in this case and in a broader context—work on dialogue and coreference has shown that we can coordinate closely and subtly with other people as we talk and listen to them, and both speakers and listeners are influenced by common ground and prior context at various levels of linguistic representation, including how they refer to things and how sentences are structured.

In terms of how language is represented, we've seen that there are certain areas of the brain that appear especially important for language, but that it is not these areas alone that contribute to our language ability. Multiple languages may be represented in different but overlapping areas of the brain, with the degree of overlap due in part to one's skill in the language. Further, bilingual representations interact during language processing and appear to have an impact on cognitive functions beyond language. Work on sign languages has begun to tell us which aspects of language are due to language itself (as opposed to the way we produce it) and as such is helping uncover what the universal underpinnings of language processing are. Current results suggest that while there are a number of similarities between spoken and sign language processing, modality does matter: there are important differences that appear to be due to form that language takes in each case.

In all of these cases, we've seen that psycholinguistic researchers have applied many different techniques and methods and that what they've found reveals that comprehending language involves not just taking in what we hear or see and processing it as it comes in, but using information from a number of sources, including prior context and previous experience, to help interpret the input and sometimes even anticipate what's coming next. Language production involves coordinating the same types of information, but with the intention to produce it. These systems appear to process these multiple sources of information in parallel when possible, including from multiple languages, and while we may not process language fully at all times, the end result is a system that is both fast and (largely) accurate, allowing us to focus on the content of what we produce or understand, rather than on how we do it.

REFERENCES

Barton, S. B., & Sanford, A. J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition, 21*, 477–487.

Birch, S., & Garnsey, S. (1995). The effect of focus on memory for words in sentences. *Journal of Memory & Language, 34*, 232–267.

Bredart, S., & Modolo, K. (1988). Moses strikes again: Focalization effect on a semantic illusion. *Acta Psychologica, 67*, 135–144.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368–407.

Christianson, K., Williams, C., Zacks, R., & Ferreira, F. (2006). Younger and older adults' good enough interpretations of garden-path sentences. *Discourse Processes, 42*, 205–238.

Erickson, T. A., & Matteson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior, 20*, 540–552.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 14–203.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11–15.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory & Language, 29*(2), 181–200.

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1366–1383.

Koh, S., Sanford, A. J., Clifton, C., Jr., & Dawydiak, E. J. (2008). Good-enough representation in plural and singular pronominal reference: Modulating the conjunction cost. In J. Gundel, & N. Hedberg (Eds.), *Interdisciplinary perspectives on reference processing* (pp. 123–139). Oxford, UK: Oxford University Press.

Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge, UK: Cambridge University Press.

Poesio, M., Sturt, P., Artstein, R., & Filik, R. (2006). Underspecification and anaphora: Theoretical issues and preliminary evidence. *Discourse Processes, 42*(2), 157–175.

Sanford, A. (2002). Context, attention and depth of processing during interpretation. *Mind and Language, 17*, 188–206.

Sanford, A., Sanford, A., Molle, J., & Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes, 42*(2), 109–130.

Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences, 6*(9), 382–396.

Sturt, P., Sanford, A., Stewart, A., & Dawydiak, E. J. (2004). Linguistic focus and good-enough representations: An application of the change-detection paradigm. *Psychonomic Bulletin & Review, 11*, 882–888.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory and Cognition, 36*, 201–216.

Traxler, M. J., Pickering, M. J., & Clifton, C., Jr. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language, 39*, 558–592.

van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language, 52*, 284–307.

van Gompel, R., Pickering, M. J., & Traxler, M. (2001). Reanalysis in sentence processing: Evidence against constraint-based and two-stage models. *Journal of Memory & Language, 45*, 225–258.

Wason, P. C., & Reich, S. S. (1979). A verbal illusion. *Quarterly Journal of Experimental Psychology, 31*(4), 591–597.

Index