

# LANGUAGE AT THE SPEED OF SIGHT STUDY GUIDE

MADISON BARKER, MOLLY FARRY-THORN, AND  
MARK SEIDENBERG

# CONTENTS

Introduction	1
<b>Part I. READING, WRITING, AND SPEECH</b>	
1. <u>The Problem and the Paradox</u>	3
2. <u>Visible Language</u>	8
3. <u>Writing: It's All Mesopotamian Cuneiform to Me</u>	14
<b>Part II. HOW WE READ</b>	
4. <u>The Eyes Have It</u>	21
5. <u>F u cn rd ths, u cn gt a gd jb n rdng rsch</u>	28
6. <u>Becoming a Reader</u>	32

# INTRODUCTION

Hi, Reader! As you've figured out---like, the title?---this is a guide to Mark's book.

I'm Mark. Who are you? You're someone interested in the book; thank you for that! You want to know about "the science of reading", and guess what? It turns out that parts of the book are very tough going for people without much background. That's not surprising! We know that comprehending a text depends on having sufficient background knowledge about the topic. Although it was nice of that reviewer from the Wall Street Journal to say that "Every teacher of young children as well as those who train them should reading this book," a major point of the book is that most educators don't have much background in the "science of reading", which then makes it a hard read.

So, we've written a guide. "We" is Mark and two able co-authors, Madison Barker (now at UC Davis) and Molly Farry-Thorn, a Washington University Ph.D. in reading.

The guide distills the main points of each chapter; it provides links to reliable, updated sources of additional information; it includes questions that can help you check your own understanding, make closer connections to your personal interests and concerns, and provide material to discuss in a book club or interest group. There's a glossary of the technical terms that will eventually include demonstrations and illustrations, not just definitions. We've linked to the studies mentioned in the book although, unfortunately, they're not all accessible for free.

We see this as a living document that will be updated frequently. Feedback and suggestions for future versions are welcome: send them to [seidenbergreading@gmail.com](mailto:seidenbergreading@gmail.com). The guide is posted on our website, [seidenbergreading.net](http://seidenbergreading.net) and can also be downloaded from there.

Thanks for your interest in the book, and for what you do!

## Inside Story

The book has three parts: Reading Writing and Speech; How We Read; The Educational Challenges. There are endnotes at the back that provide sources for quotations and other material, and pointers to additional resources, including URLs for videos and helpful websites. You can find the endnotes with live, clickable links on [seidenbergreading.net](http://seidenbergreading.net), along with demonstrations of various phenomena, simple versions of experiments, a glossary, and other useful stuff.

PART I

READING, WRITING, AND  
SPEECH

# 1.

## THE PROBLEM AND THE PARADOX

---

### OVERVIEW

- This chapter explains what the book is about: the science of reading and its relevance to educational practice and improving literacy.
- Reading is one of the highest achievements of human intelligence. Many other species communicate but none have anything like reading and writing.
- People are very good at reading without knowing much about how it works. Our intuitions about how we read don't go very deep because most of the activity occurs without conscious awareness. We are aware of the result of reading, not how it happened. It takes systematic research to unpack what is going on "under the hood".
- The "Science of Reading" is a body of basic research about how reading works and the capacities it depends on: language, learning, memory, vision, thinking and others. Much is known about skilled reading, learning to read, conditions that promote or interfere with reading, how reading works in different languages and writing systems, and other topics. This research has important educational implications.
- The *problem* is that literacy levels in the US are too low: a large proportion of children and adults have only basic reading (and writing) skills. The *paradox* is: if the science of reading has been so successful, why are there so many poor readers?
- One answer is that reading science has had little impact on educational policies and practices. Teachers (and administrators and the people who develop curricula and instructional materials) aren't exposed to this research as part of their professional training; what they are told about how reading works and how children learn is limited and out of date. Making better use of this science might allow more teachers to be more effective and more children to succeed.
- The book has two purposes: first, to describe this science, how it is conducted, and what has been learned; second, to examine the disconnection between this science and educational policies and practices, and how it might be overcome. The goal is to allow more people to participate in debates about how reading should be taught, and work toward practices that will yield better outcomes.

## tl;dr

(too long, didn't read)

Reading is important; it takes science to understand how it works; the science of reading can help to improve literacy if more people understand it and connect it to effective educational practices.

## CHAPTER SECTIONS

1. Introduction – p. 3
2. Sputnik Lands on USA – p. 5
3. A Science of Reading – p. 12

### Introduction

Reading is studied by psychologists, psycholinguists, and cognitive neuroscientists around the world. This body of research is now called reading science, which is a scientific enterprise that endeavors to understand the many cognitive abilities that are involved in reading – including, but limited to, our visual, language, emotional, decision-making, and memory systems. Although reading is something you do every day, most of what you do when reading is subconscious and therefore requires scientific study to be understood. Intuitions and personal observations are not sufficient. They are also highly inconsistent and often contradictory. From the science we've learned a lot about how reading develops, the mental and neural operations that support skilled reading, and how they can be impaired to cause reading impairment. Much of this knowledge has wide consensus in scientific circles.

### Sputnik Lands on USA

National and international assessments of reading have created valid concerns about literacy levels in the US. Poor reading has numerous negative consequences: difficulties with reading have downstream effects on learning math and every other subject. In adults it is associated with lower incomes, poorer health, higher incarceration rates, and shorter lifespan. Low literacy has numerous causes. Poverty is one, and it is sometimes used as an excuse for poor reading outcomes (see Chapter 10). Low literacy isn't limited to people with low socioeconomic status (SES); the US trails countries such as Korea and Canada at every income level. Moreover, quality education is even more important for people from low SES backgrounds, who have less access to other resources such as tutors, computers, and the Internet. Traditional literacy is also being affected by cultural changes such as the availability of smartphones that don't require high reading or spelling skills. Some factors that affect literacy are hard to change. The situation is far

from hopeless, however. There are things that can be changed that would have enormous impact. One is how reading is taught. The methods that are commonly used to teach children to read are not in line with research findings. Because of cultural differences between the two fields, there is a disconnect between the science of reading and educational practice. Bringing the basic science of reading into the field of education and into the classroom could improve outcomes for many children.

## A Science of Reading

A science of reading has emerged from psychology and related disciplines over the past century. Why study reading? First, because it is a uniquely human activity. Studying it reveals things that are essential about being human and distinguish us from other species. Second, because of the importance of reading in modern life. People who are skilled readers have advantages over those who are not. If we understand how reading works, we can help more people gain this advantage. Finally, reading is a way to study many aspects of human behavior and its brain bases. Reading is a relatively recent technology that humans created out of existing cognitive and linguistic parts. So, we might study reading for its own sake or because it is a tool for understanding all the capacities it relies on.

## STUDIES MENTIONED

Page 5: [Program for International Student Assessment \(PISA\)](#)

Page 6: [National Assessment of Adult Literacy](#)

Page 6: [National Assessment of Educational Progress \(NAEP\)](#)

Page 9: [Hayes, Wolfer, & Wolfe](#) (1996)

## KEY TERMS

**Magnetic resonance imaging machine** (p. 4) – a technology that produces detailed images of the brain or other organs and tissues in the body.

**Sputnik moment** (p. 5) – the point when a country or society realizes that considerable effort is needed to catch up to other countries. In the case of reading, it was occasioned by the release of the PISA assessments of academic performance.

**Program for International Student Assessment (PISA)** (p. 5) – worldwide study that evaluates 15-year-old students in reading, mathematics, and science.

**Organization for Economic Co-operation and Development (OECD)** (p. 5) – an intergovernmental organization, comprised of 37 member countries, founded to stimulate economic progress and world trade.

**National Assessment of Educational Progress (NAEP)** (p. 6) – congressionally mandated assessment of United States students in various subjects.

**Socioeconomic status (SES)** (p. 8) - the social standing or class of an individual or group. It is often measured as a combination of education, income, and occupation.

**Deep dyslexia** (p. 10) – a form of acquired dyslexia that results in semantic errors (e.g., city read as town) during reading and the impairment of nonword reading.

**Educational neuroscience** (p. 10) – an emerging field that unites neuroscience research with educational research to study the brain bases of skills such as reading and math.

**Literacy** (p. 10) – the ability to read or write, some researchers insist that these abilities must be considered in a social and cultural context. In some contexts, it refers to the exchange of information by both linguistic and nonlinguistic means (including visual, digital, and technological literacy).

## RELATED CONTENT

[Nation's Report Card](#) - p. 7

[Transcranial Magnetic Stimulation and Reading](#) - p. 10

[Alex the African Grey Parrot](#) - p. 12

## DISCUSSION QUESTIONS

1. Describe three consequences of children's reading failures. What are their implications for the child, family, community, and beyond? [p. 7]
2. Mark asserts that there is a disconnection between the science of reading and educational practice. That's pretty harsh! Is this claim consistent with your own experience? [p. 9]
3. Do you think it's realistic to think that educators (not just teachers but principals, superintendents, the professors who teach them and create curricula) could gain sufficient expertise in the science of reading to make use of it? [p. 9]

4. Educational research often focuses on “literacy” instead of “reading”. When people use one term or the other, what are they trying to signal? [p. 10]
5. What are the basic tools you need to be an advocate for change in reading instruction? [p. 11]
6. "Reading is a technology": what does that mean? Language is not a technology: why not? Is math a technology? What about fire? Discuss! [p. 12]

## ACTIVITY

Before you go any further, create a file on a suitable device where you can keep questions, notes, ideas. This can be your first entry: In a paragraph or so, describe your personal theory about how people read—how YOU read. What does it take to be a skilled reader? Just list the basic ideas and keep the document for later. Over the course of reading and discussing this book you will be exposed to research about the mental and neural mechanisms and types of knowledge that underlie skilled reading. Later we'll ask you to do this exercise—describing your theory of reading—again. It should be interesting to see how the two do and do not overlap.

## HANDY QUOTES

"Being an expert reader doesn't make you an expert about reading. That is why there is a science of reading: to understand this complex skill at levels that intuition cannot easily penetrate."

"A look at the basic science suggests specific ways to promote reading success. These do not require more testing or new federal laws; they do not require vast infusions of money; they are not based on classroom computers that treat learning like a video game or other faddish uses of technology. What they require is changing the culture of education from one based on beliefs to one based on facts."

## INSIDE STORY

This quote, "...on the Internet nobody knows your website is a dog," is a pun on this famous New Yorker cartoon. Do people still use the expression "what a dog!" to mean something terrible or ugly?

## 2.

# VISIBLE LANGUAGE

---

## OVERVIEW

- The topic of this chapter is the relationship between reading and speech. They are very similar and closely connected but differ in important ways that affect how they are learned and used.
- Terms: We use "reading" and "written language" interchangeably. The terms refer to both reading itself (comprehending print) and writing (producing it). Similarly, "spoken language" and "speech" are also interchangeable; they refer to both listening (comprehending spoken utterances) and talking (producing them).
- Reading depends on speech. Speech existed long before reading; writing was invented as a way to represent spoken language in a more durable form. Children repeat this pattern: they learn to speak before they learn to read and write.
- Most messages can be communicated using either speech or reading, but many things can be expressed more easily in one than in the other.
- Writing systems represent information about the sounds and meanings of words. However, they leave out many properties of speech that affect understanding. Examples include which syllable or words are stressed; prosody (how phrases are pronounced); and cues to the speaker's emotional state (such as anger or glee) and identity (gender, age, race/ethnicity).
- We use visual embellishments such as emojis, underlining, *italics*, and CAPITAL LETTERS to restore some of this missing information (for example, to convey sarcastic or humorous intent). These tools aren't very precise or used very consistently, however.
- Aside from its greater permanence, writing's main advantage is that it allows more complex linguistic communication than speech (because texts can be re-read and re-written).
- The definition of learning to read is learning how the spoken language a child already knows is represented in a new code, writing. Everything the beginning reader knows about language derives from their prior experience with speech. Reading requires being

able to access that knowledge. The way to do that is by taking advantage of the fact that writing systems *represent* spoken language.

- For skilled readers, the written and spoken codes become deeply intertwined. Literacy affects how the brain represents and processes speech; spoken language is deeply implicated in learning to read and in skilled reading. Children whose connections between print and speech are weak (because of inadequate instruction or other conditions) struggle with reading.
- Learning a spoken language and learning to read are vastly different. Children rapidly pick up spoken language through exposure and use. The capacities that make this possible evolved over many eons in human history. Reading, a recent technology based on speech, involves other skills such as learning arbitrary visual elements (such as letters), how they combine to represent words, and their associations with sound and meaning. Whereas learning to talk is natural, learning to read is an "unnatural act" (Phil Gough) that almost always requires instruction and guidance.

### tl;dr

Reading depends on speech. Writing systems are codes for representing the sounds and meanings of words; learning to read is the process of learning how. As reading skill develops, the spoken and written codes become so deeply intertwined they are inseparable. Reading isn't just about spelling and meaning; it is also about sound, intrinsically and unavoidably.

## CHAPTER SECTIONS

1. Introduction - p. 15
2. Each of These Things is Not Like the Other - p. 16
3. Mashup! - p. 19
4. Hear My Words - p. 20
5. Just Cheat - p. 24
6. What Is It Like to Be the Word "Bat"? - p. 26

### Introduction

Reading and speech are very similar but differ in important ways. Reading is secondary to (= dependent on, derived from) spoken language. Speech evolved in the species; writing was invented much later. This pattern is repeated in child development: learning a spoken (or signed) language always precedes reading. The beginning reader's initial challenge to learn how a writing system (such as the English alphabet) represents the sounds of words. Making this

connection allows the child to access their knowledge of spoken language and everything they've learned through their use of it. Connecting print and sound can be difficult because the spoken and written codes have different properties. Moreover, in English the correspondences between spelling and sound are complex, with many seeming inconsistencies. Difficulties in connecting print and speech are common among struggling readers and a prominent characteristic of dyslexia.

### **Each of These Things is Not Like the Other**

Reading and speech are alternative ways of using language, but the relationship between them is complicated, with numerous similarities and differences. Many of the differences relate to the fact that speech is fast-fading (the sounds rapidly fade away), whereas text is not. The ways they are learned are also very different. Children acquire their first, spoken language through exposure and interactions with other speakers. Parents do not teach children to talk; they provide experiences that enable them to learn. Learning to read on the other hand, with few exceptions, requires instruction, practice, and feedback.

### **Mashup!**

The differences between spoken and written language arise from the properties of the auditory and visual modalities, which employ different sensory systems (hearing vs. seeing) and motor systems (the mechanisms for producing speech vs. hand movements as in writing or typing). Texting is interesting because it combines properties of speech and reading. It fulfills a narrow function: rapidly producing simple messages on devices with minimal keyboards such as smartphones. Usually, combining features of print and speech creates an artificial code that is difficult to learn and use.

### **Hear My Words**

Each way of using language has advantages and disadvantages. Advantages of writing include the fact that word boundaries are clearly indicated by the use of spaces and homophones such as pair/pare/pear can be differentiated through spelling. However, writing is efficient to produce and comprehend because it leaves out information that readers can (usually) supply on their own. For example, English doesn't indicate syllabic stress (the difference between PERmit and perMIT) because we can figure it out from our knowledge of spoken language and the context in which a word appears. The system does not work perfectly, but errors are usually minor and easily corrected. The alternative—representing stress in order to avoid these minor errors—would make texts harder to read because there would be more symbols to process. The costs would greatly outweigh the benefits.

## Just Cheat

Authors can use typographic embellishments, such as bold face or italics, to clarify the intended meaning of a text. They are mainly used to restore missing information about pronunciation that affects meaning, such as which word gets heavily stressed in the sentence "Fred didn't take the text yesterday". The use of such tricks differs across languages and cultures; they are used less often in writing Mandarin, for example. Even within English, only some properties of spoken language can be conveyed this way. Heavy stress on individual words is easy to convey through typography, whereas prosody is much harder to represent.

## What Is It Like to Be the Word "Bat"?

Learning to read changes our perception of spoken language. The word written BAT is spoken as a single syllable "bat". We think of it as consisting of three discrete parts: the initial consonant "b", the vowel "a" and the final consonant "t". These parts are called phonemes. Phonemes are tricky. Spoken words don't consist of sequences of phonemes. A recording of the syllable "bat" cannot be cut into three parts corresponding to three discrete sounds. Phonemes are an abstraction: we treat spoken words as if they consist of phonemes because it is crucial for reading. Alphabets represent the sounds of words at the level of phonemes. Treating "bat" as though it consists of three discrete phonemes makes speech more like writing. Units in the spelling of a word now correspond to units in its pronunciation. That makes learning the correspondences between print and sound much easier.

So, phonemes are like an alphabet for speech. The famous "Alphabetic principle" is that units in the written code (graphemes, which are single letters such as B and letter combinations such as SH) correspond to units in the spoken code (phonemes). "Phonemic awareness" refers to treating spoken words as consisting of phonemes. Phonemes begin to emerge through the use of spoken language (e.g., rhyming games), but it takes exposure to print to finish the transition to treating speech phonemically.

These are important concepts that will be revisited throughout this book. If you want to learn more about them now, see the suggested activity below.

## STUDIES MENTIONED

Page 17: [\*Goldin-Meadow & Mayberry\*](#) (2001)

Page 17: [\*Marlsen-Wilson\*](#) (1975)

Page 17: [\*Gernsbacher, Varner, & Faust\*](#) (1990)

Page 19: [\*Montag, Jones, & Smith\*](#) (2015)

Page 23: [Bentin & Frost](#) (1987)

Page 23: [Seidenberg & McClelland](#) (1989)

Page 27: [Lieberman, Shankweiler, & Liberman](#) (1989)

Page 28: [Morais et al.](#) (1979)

Page 29: [Seidenberg & Tanenhaus](#) (1979)

## KEY TERMS

**Natural language** (p. 15) – Spoken and signed languages that are humans' primary mode of interpersonal communication. NLS evolved among communities of users without conscious planning or premeditation. Excludes written language (a technology for representing spoken language) and many other language-like systems (e.g., computer languages; animal languages).

**Dyslexia** (p. 16) – childhood condition that interferes with learning to read. Causes and severity vary; typically manifested as difficulties learning about print and its connection to spoken language.

**Homophones** (p. 20) – A word that is pronounced the same as another word but has a different meaning. The two words may be spelled the same (homonym) or differently (heterograph). In English, examples are WATCH, ROSE, HOARSE/HORSE, and TWO/TO/TOO.

**Orthography** (p. 22) – the written code used to represent a language. English uses an alphabetic orthography.

**Phonology** (p. 22) – the sound patterns of a particular language.

**Syllable** (p. 22) – a unit of pronunciation having one vowel sound, with or without surrounding consonants, forming the whole or a part of a word.

**Quasiregularity** (p. 23) – systems that have rule-like patterns but also admit patterns that deviate from them. Written English is quasiregular. There are rule-like patterns (SAVE GAVE PAVE) and exceptions (HAVE). Note that the "exceptions" (or "sight words") are not arbitrary: they partially overlap with the rule-governed cases. The past tense is also quasiregular: bake-baked, rake-raked, take TOOK.

**Stress deafness** (p. 24) – difficulty exhibited by second language learners of mastering stress patterns for words (e.g., "orbit" has primary stress on the first syllable, but "obese" has primary stress on the second).

**Prosody** (p. 25) – sound patterns that occur across sequences of words. In English, primarily conveyed by intonation (rising and falling pitch) and stress (how heavily a word is emphasized).

**Phoneme** (p. 26) – the vocal gestures from which words are constructed in a language; the smallest unit of speech that serves to distinguish one utterance from another.

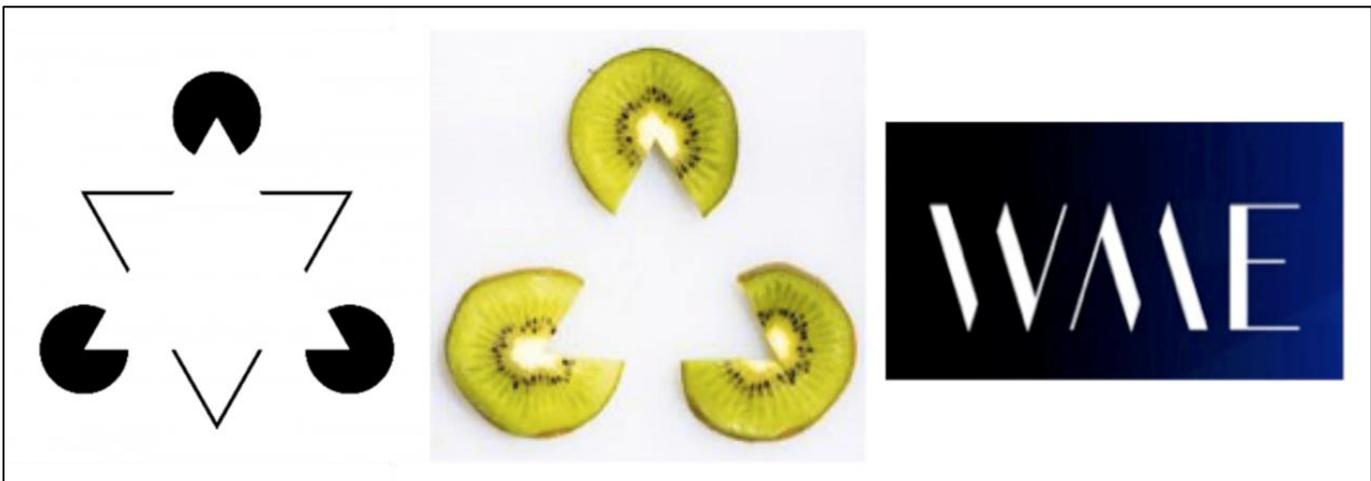
**Articulatory gestures** (p. 27) – the actions necessary to enunciate language (e.g., mouth movements in speech and hand movements in sign language).

**Graphemes** (p. 27) – The term "letter" refers to the 26 symbols in the alphabet. "Grapheme" refers to a single letter or combination of letters that corresponds to a single phoneme. P and PH are graphemes that represent phonemes ("p" and "f", respectively).

## RELATED CONTENT

Page 20: [Signed English vs ASL](#). Signed English is slower and less efficient.

Page 22: "We fill in missing information all the time." As in many visual illusions:



Left: A famous visual illusion that illustrates how we fill in missing information. The triangle that faces up is seen even though it has no lines; the one that faces down is seen even though the lines are interrupted by white space. More information [here](#).

Middle: A witty version of the Kanizsa triangle by the artist Yve Moreau (h/t Mindhacks).

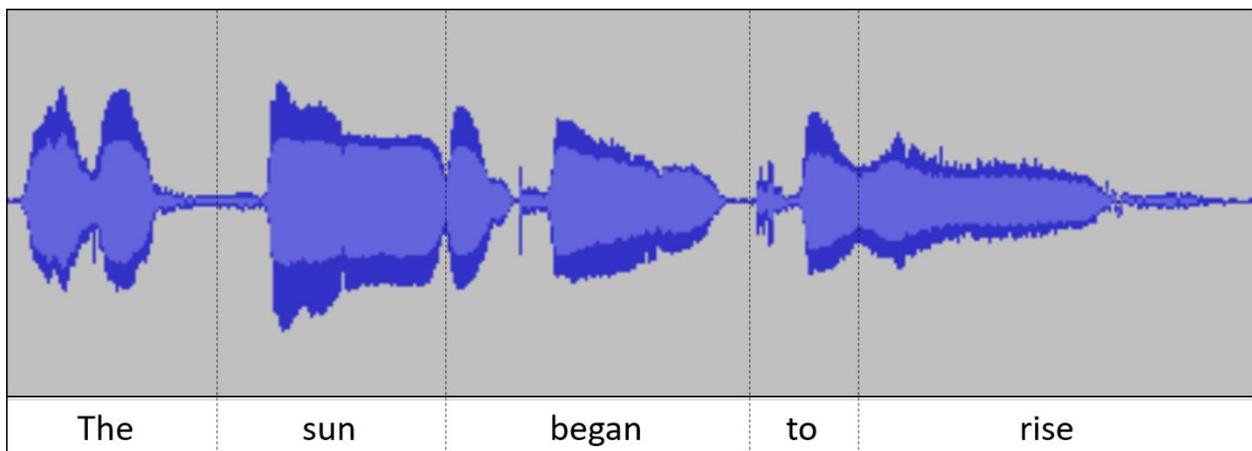
Right: Incomplete fonts are readable even though parts of letters are left out. This is the logo of the William Morris Endeavor agency. Some lines in the W and M are only implied.

Page 23: Goodnight Moon in Hebrew. In written Hebrew, the vowels are only included in books for beginning readers. Skilled readers can read with the vowels deleted: they fill in the missing information. Including the vowels slows them down because the information is redundant (= superfluous = TMI).



Page 26: [Russian interpretation](#) of Ronald Reagan's bombing remark.

Page 27: Discreteness of words as byproduct of written language. In speech there are no "spaces" between words, as can be seen in the sentence waveform below. (The dotted lines only indicate the approximate locations of transitions from one word to the next.) See and hear another demo on Mark's [website](#) (the first video under the Chapter 2 heading).



Page 27: [Sesame Street spelling](#) illustrating that spoken words really don't consist of a sequence of discrete sounds (phonemes).

Page 28: You can try the rhyming experiment [here](#) (it's the third video under the Chapter 2 heading).

## DISCUSSION QUESTIONS

1. The text describes ways in which speech and reading differ. For example, speech is "fast-fading". Can you think of exceptions to these assertions? Are there cases where reading is more like speech, or vice versa? [p. 18]
2. Texting is a 'mashup' of written and spoken language. Texting conventions such as abbreviations (BRB, GTGN) and emoticons (such as :) and : D) were very popular in the flip phone era (1990s), when typing a message was awkward. Now that smartphones with full keyboards and predictive text have taken over, do people still use texting tricks? For what, and why? [p. 19]
3. Audiobooks present texts that were written for one medium (visual, reading) in another medium (auditory, spoken). We've said that texts and speech have different properties; do these differences have any impact on the experience of "reading" an audio book vs. reading the corresponding text? Are they equally easy to comprehend? Is the choice between them just a matter of convenience or personal preference? What kinds of books work better or worse in the audio format?
4. Can you remember any cases where you a text you wrote was misunderstood because properties of spoken language weren't represented? [p. 21]
5. Why is Hebrew a good example of written language being efficient? [p. 23]
6. What does it mean to say that phonemes are like "spoken letters"? How are phonemes similar to letters? How are they different? [p. 27]

## ACTIVITY

If you're still unsure about what phonemes are, take a look at [this](#). It should help a lot. Main point: sometimes the number of letters in a word corresponds to the number of phonemes (e.g., CAST, STRAP), but often it does not (ACHE, DOUGH, STRAIT).

## HANDY QUOTES

Writing and speech are "like a couple of linguistic codependents with serious boundary issues".

"For people who can read, there are no pure representations of the sounds of words in the brain because they've been contaminated by spelling."

## INSIDE STORY

Who is John Cleese (p. 21)? Comedic actor known for [Monty Python](#) and [Fawlty Towers](#). The grasshopper joke turns on the fact that there was (is?) a popular cocktail called a [Grasshopper](#).

Who the heck is Clive James (p. 24)? A famous TV presenter and humorist in [Great Britain](#) and Australia; also a poet who translated Dante's Divine Comedy. This would be like Stephen Colbert translating Marcel Proust from the original French. I shouldn't have used such an obscure celebrity, but the example was a gem.

"What is it like to be the word 'bat'?" is a pun on a famous paper by the philosopher Thomas Nagel, "What is it like to be a bat?"

### 3.

## WRITING: IT'S ALL MESOPOTAMIAN CUNEIFORM TO ME

---

### OVERVIEW

- What is a writing system, how do writing systems work, and how were they invented? Most important: do the properties of writing systems affect how people read? Or do we all read in essentially the same way?
- Writing is one of the critical inventions in human history. The development of modern civilization could not have occurred without the massive increase in the creation, retention and transmission of information that writing enabled. So how did it happen?
- Humans used pictures to communicate for thousands of years before writing came along. A picture of an owl, for example, represented an owl, the thing itself. This is depiction. Writing came into existence with the insight that graphical elements (pictures, abstract shapes, squiggles) could instead be used to represent language (e.g., the image of an owl for the spoken word 'owl').
- We don't know who invented writing, but it probably happened independently in 3 locations: Mesopotamia (modern Middle East), China, and Mesoamerica (modern Central America/Southern Mexico).
- The token theory is a plausible account of how writing emerged in Mesopotamia. It came about gradually rather than through a stroke of genius.
- After the initial development of primitive writing systems, it took several thousand additional years to develop ones that could fully represent a spoken language using a manageable number of graphical elements.
- Despite their apparent differences, all modern writing systems work essentially the same way: the graphical elements provide clues to the sounds and meanings of words. How people read is very similar as well. The neural systems for reading are the same with some writing-specific adjustments at the "front end", in parts of the visual system involved in recognizing the symbols.

- Writing systems differ in what the graphical elements represent (phonemes, syllables, other phonological units, cues to meaning). Units in a writing system have to be congruent with units in the spoken language they represent. A syllabary (symbols represent syllables) works for Japanese but not Chinese or English; an alphabet (symbols represent phonemes) works for English but not as well for Japanese or Chinese. There isn't any "best" writing system because one size does not fit all.
- Many of the most important advances in the development of writing came from mismatches that arose when a writing system that worked for one type of spoken language was borrowed for use with a very different type of spoken language. Trying to write Greek, for example, with the Phoenician alphabet (which had no symbols for vowels) was unworkable and led to the development of full alphabets like the one for English.

## tl;dr

Writing systems are ways to represent spoken language. Figuring out that visual images (pictures, scribbles) could be used to represent speech was a huge breakthrough. All writing systems use symbols that provide clues to the sounds and meanings of words. How they do this is determined by properties of the spoken language they represent. Skilled reading works essentially the same way across language and writing systems: the reader uses knowledge of how a writing system represents sound and meaning.

## CHAPTER SECTIONS

1. Introduction - p. 15
2. Each of These Things is Not Like the Other - p. 16
3. Mashup! - p. 19
4. Hear My Words - p. 20
5. Just Cheat - p. 24
6. What Is It Like to Be the Word "Bat"? - p. 26

### Introduction

The invention of writing shaped human civilization as we know it. How did writing systems come to be? We do not have an exact record of what happened, but there are many clues from ancient artifacts.

### What the History of Writing Has to Say About Reading

How writing came into existence is a fascinating question, but the history of which writing

systems succeeded and failed provides unique evidence about how writing and reading work. The development of writing involved four major advances: (a) using graphical elements (pictures but also abstract shapes and scribbles) to represent spoken language; (b) being able to represent an entire language using a manageable set of basic elements; (c) discovering that spoken words consist of phonological components (such as phonemes, syllables), which writing systems represent; (d) achieving alignment between properties of writing systems and the spoken languages they represent.

### What Ought to Be in Pictures?

Humans were creating pictures for about 30,000 years before the invention of writing. Pictures (depictive images that resemble their subjects) are fascinating, but their communicative range is limited compared to language. Writing came about when pictures (and other graphical elements) were used in a new way: as symbols for representing language. Example:

A pictograph of an owl was initially used to depict an owl. Later (in Egyptian hieroglyphic writing) it was used to represent the spoken word for owl, and then also the first sound in the word. So:



A pictograph. When used to represent an owl, an example of depiction. When used to represent the spoken word for owl it is a linguistic symbol. Symbols don't have to resemble what they represent, and pictographs gave way to arbitrary graphical elements such as letters.

The token theory: A plausible theory (by Schmandt-Besserat) of the origins of writing in Mesopotamia. In her eyes it was an accidental by-product of something else.

1. People in Sumeria and other areas made a wide assortment of tokens, small clay objects in abstract shapes. The tokens—a forerunner of money—were used in trading barley, cows, oil, and other commodities. They were symbolic objects: they did not look like the things they represented.

2. The tokens were widely used; thousands of them have been recovered. Clay vessels—containers—were created to store tokens. To keep track of which type of token was



Tokens from about 3300 BC. They include ones for one ram (sheep), a quantity of oil, a quantity of wheat, a length of rope, and others. Note that they do not resemble these things: they are abstract.

stored inside, a label was created by making an imprint of the token on the outside. Thus a 3-D token was turned into a 2-D image.

3. That label was a symbol for the tokens inside, which were in turn symbols for objects such as cows. Writing began when these visual labels started to be associated with the spoken words for cow, oil, etc. Proto-cuneiform—the earliest known writing—writing incorporated many of these symbols.

The symbols were not pictures because the tokens they are based on look nothing like the things they stand for. Thus, the earliest writing did not arise from pictures. Writing used abstract symbols from the start. (Pictographs were included later in some systems, to create graphical symbols for a larger number of words.)

### **A Picture Is Worth How Many Words?**

Languages allow an infinite number of messages to be expressed using a finite set of elements: spoken words and the parts they are composed of. A writing system has to do the same thing, using graphical symbols. Following the initial breakthrough with proto-cuneiform, it took another 2000 years to develop writing systems capable of representing entire languages. Languages contain too many words to use a different graphical symbol for each one. The Sumerians got around this by using their graphical elements for multiple purposes. For example, a symbol might represent two unrelated words. This gave them a way to represent abstract words. The words LIFE and ARROW were pronounced the same in Sumerian. The symbol for the word ARROW was also used for the word LIFE. Sumerians had discovered homonyms: words with the same spelling and pronunciation but unrelated meanings (in English, words like WATCH, TIRE, BANK). They also figured out how to combine symbols to form new words, a basic principle used in all subsequent writing systems. A word might be represented by a symbol that provided a cue to its meaning plus another that provided a cue to its sound. Combining these types of cues has been called the “charade principle,” after the game charades. The graphical symbols in cuneiform were eventually used so inconsistently for so many different purposes that the system became excessively complicated to use or pass on to others. The Sumerians nonetheless get credit for developing the first full writing system: anything that could be said in Sumerian could be written in the later, advanced forms of cuneiform.

### **Writing Is for Sound and Meaning**

The graphical symbols in cuneiform served multiple functions, which made them hard to interpret. A symbol might be a cue to something about the sound or meaning of a word, but it was hard to tell exactly what. Nonetheless the idea of combining sound and meaning cues was a

key advance. In modern written Chinese, about 80 percent of the characters incorporate cues to sound and meaning. Japanese took this strategy a step further, creating separate types of graphical elements to represent phonology and semantics. Modern Semitic languages such as Hebrew and Arabic convey sound and meaning cues in yet another way. Most words have triconsonantal roots (like K-T-B, but in their letters). Words that share this root are usually related in meaning. In alphabetic writing systems, the symbols represent sounds (phonemes) but groups of phonemes form morphemes, which are meaning-bearing units like the BOOK in BOOK/BOOKS/BOOKING or the past tense morpheme -ED. In this way alphabets too convey cues to both sound and meaning.

### Finding Phonemo

The words in spoken language are composed of units that are recombined. Languages differ in which units are the primary building blocks. Writing systems work when units in the written code represent the primary units in the spoken language but this was only discovered through trial and error over many centuries. As the use of writing spread, a writing system for one language would often be imported for use with another language. A variety of unrelated languages were spoken in the Middle East where writing developed. Cuneiform was effective for Sumerian, a language with mostly monosyllabic words. The symbol for the word SHEEP also represented the syllable in its pronunciation. Using symbols to represent syllables did not work for other nearby languages, however. The mismatches that occurred when a writing system for one type of language was adopted for a different type of language led to many innovations. For alphabetic writing, the important breakthrough was learning to treat words as consisting of parts we call phonemes. This came about through a series of mismatches between writing systems and languages, with the crucial innovation—symbols to represent vowels—occurring in Greece. Children learning to read an alphabet repeat this process of discovering phonemes.

Is there a best writing system? No, because languages differ in ways that affect which units need to be represented in writing. To succeed, a writing system needs to represent a spoken language at the level of detail appropriate for that particular language.

### STUDIES MENTIONED

Page 33: [Schmandt-Besserat](#) (1992)

Page 38: [Gough & Hillinger](#) (1980)

Page 38: [Frith & Frith](#) (2005)

Page 42: [Schmandt-Besserat](#) (1986)

Page 49: [Dingemanse et al.](#) (2015)

## KEY TERMS

**Pictograph** (p. 36) – a standardized picture that acts as a symbol for something (a word, a phoneme or syllable, a semantic category, etc.).

**Graphical element** (p. 36) – in this context, a written character/symbol that represents language.

**Theory of mind** (p. 38) – the capacity to attribute mental states to oneself and to others.

**Token** (p. 41) – in this context, a graphical element which is used first as a label for an object in the world and then shifts to being used to abstractly represent a word.

**Lone boffin theory** (p. 41) – writing system was created by a single, unknown individual.

**Proto-cuneiform** (p. 41) – precursor to written language, an early writing system consisting of pictographs, rather than a written representation of spoken language.

**Morphemes** (p. 44) – the smallest unit of meaning in speech. Some morphemes are not words, but they still convey meaning (e.g., in cats the morpheme -s conveys that there multiple cats).

**Logograph** (p. 44) – symbols in a writing system that represent whole words.

**Homonyms** (p. 45) – words that are spelled alike or sound alike but have different meanings.

**Rebus principle** (p. 45) – using pictographs to represent the phonology, rather than semantics, to convey meaning.

**Polyphony** (p. 45) – using one sign to represent several semantically related words, all pronounced differently.

**Semantics** (p. 46) – the meaning of a word, phrase, sentence, or text.

**Charade Principle** (p. 46) – representing words with two signs, one providing the reader with a cue to the meaning and the other providing a cue to the sound.

**Phonological awareness** (p. 50) – general term for knowledge of the phonological properties of language, including components of words such as syllables; similarities between words, such as rhyming; recognition and use of prosody, and others. Contrasts with **Phonemic awareness** – ability to treat words as though they consist of discrete phonemes.

## RELATED CONTENT

Ideographs, pictographs, logographs: what are they? Everything you need to know: [here](#).

Page 32: [Online encyclopedia of writing systems and languages](#)

Page 32: "Exotic writing systems inscribed on their bodies..."

I got this tattoo 17 years ago in NYC. Tattoos were not totally legal in the city yet and I was underage. Internet was not in full swing either. The tattoos were supposed to have read, "Strength and Courage." I'm sure they don't. LoL! If you could tell me what the characters mean (if anything), I would appreciate it.

Thanks,

JKreasy



小畜 "little animal" and 大過 "big mistake".

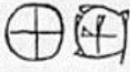
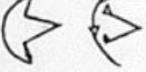
Many tragic/entertaining examples [here](#).

Page 36: [Cave paintings](#)

Page 37: [Timeline of writing systems](#)

Page 38: [Clay tokens](#)

Pages 44-46: Tokens and cuneiform symbols derived from them.

					Sheep
					Cattle

Tokens (left) do not represent what they refer to. The graphical elements in early cuneiform closely resembled these tokens. In later cuneiform (other boxes) the symbols no longer looked like the tokens.

Page 48: Combination of sound + meaning cues in Chinese writing system



# 雞 Rooster

This character is formed by the phonetic word on the left side ('hai') and the radical on the right meaning 'bird'. The character for rooster is pronounced 'gai'.

www.blss.portsmouth.sch.uk

Simple example of a word consisting of phonetic + radical (cues to sound and meaning).

Page 49: [K-T-B words in Hebrew and Arabic](#)

### Standard Japanese

すべての人間は、生まれながらにして自由であり、かつ、尊厳と権利とについて平等である。人間は、理性と良心とを授けられており、互いに同胞の精神をもって行動しなければならない。

### With Furigana appended:

すべての人間は、生まれながらにして自由であり、かつ、尊厳と権利とについて平等である。人間は、理性と良心とを授けられており、互いに同胞の精神をもって行動しなければならない。

### Written in Katakana:

スベテノニンゲンハ、ウマレナガラニシテジユウデアリ、カツ、ソングントケンリトニツイテビョウドーデアル。ニンゲンハ、リセイトリョーシントヨサズケラレテオリ、アガイニドーホーノセイシンヨモツテコードーシナケレバナラナイ。

### Written in Hiragana:

すべてのにんげんは、うまれながらにしてじゆうであり、かつ、そんげんとけんりについてびょうどうである。にんげんは、りせいとりょうしんとをさずけられており、たがいにどうほうのせいしんをもってこうどうしなければならない。

**Another look at Japanese writing.** Standard Japanese: sentence illustrates the mixing of Kanji and two types of Kana scripts. Furigana are symbols that provide clues to Kanji pronunciations; helpful for learners. The sentence can be written in both types of Kana, but even without knowing Japanese you can probably tell that finding the boundaries between words is harder than when the scripts are mixed.

## ラドクリフ、マラソン五輪代表に1万m出場にも含み

Kanji, red; hiragana, blue; katakana, green; Latin letter and number, black

*Radokurifu, Marason gorin daihyō ni, ichi-man mētoru shutsujō ni mo fukumi*

Same sentence written in Latin alphabet (romaji)

“Radcliffe hints she may compete in Olympic 10,000 m as well as marathon”

Adapted from [https://en.wikipedia.org/wiki/Japanese\\_writing\\_system](https://en.wikipedia.org/wiki/Japanese_writing_system)

Mixing of scripts in Japanese. Color is used to identify symbols in the several scripts used in writing a simple sentence. (For illustration only; color is not used this way in actual texts.) Kanji are characters like those in Chinese; the two Kana scripts represent spoken syllables. Every sentence could be written exclusively in Kana, and also in Romaji, the Latin alphabet, but those are harder to read than the mixed Kana-Kanji script.

## DISCUSSION QUESTIONS

1. Language evolved in the spoken modality (talking and listening) and writing only came around much later. Why didn't the opposite happen: use of written language, followed by invention of speaking? Why can't children learn to write before they learn to talk?
2. What are the four developments that were critical to the development of modern writing systems? [p. 36]
3. What was the role of pictures/pictographs in the development of writing? [p. 39]
4. In the Schmandt-Besseret theory, language is a by-product of other cultural innovations. Explain. [p. 42]
5. Describe one of the mismatches in the history of writing that led to an important new insight. [p. 44]
6. Are alphabets the most advanced form of writing? [p. 52-53]

## ACTIVITY

Could you create a writing system out of emojis? *The dog chased the cat around the house.* Try to write this sentence using emojis (get help [here](#)). How far did you get? Now imagine if there was an emoji for every word in the sentence. That would be a logographic form of writing: each symbol represents a word. The book says that no writing system can be completely logographic: why not?

## HANDY QUOTES

“The stories about where and when writing was invented are what I would call honest confabulations...docudramas of the academic world, “based on a true story.”

“The astonishing fact is that tens of thousands of years elapsed between depictive drawing and the emergence of writing. What *took* so long? Using graphical elements to represent language rather than the world.”

“The characteristics of writing systems are determined by the ways they represent spoken languages, and the ways spoken languages, in turn, represent meaning. Writing systems are alike because they represent phonology and semantics, though the solutions vary in detail. This property has an important implication for how we read: reading is not just about spelling; it is inherently also about phonology and semantics because that is what writing systems represent.

PART II

# HOW WE READ

## 4.

# THE EYES HAVE IT

---

## OVERVIEW

- Reading is the comprehension of language in visual form: “language at the speed of sight”. Reading isn’t a single activity: it encompasses everything from recognizing individual words to skimming texts to poring over complex texts. Descriptions of how reading works usually focus on reading a text carefully with good comprehension.
- Properties of human vision have a huge impact on how we read. The visual system evolved for other purposes: recognizing faces and objects, and perceiving motion, distance, space, and other properties of the visual world. Reading repurposes this system for a very different activity, recognizing two-dimensional sequences of abstract symbols.
- If you are a good reader, texts can be comprehended more rapidly than speech. Listeners are at the mercy of speakers who control how rapidly they talk and how clearly ideas are expressed. Reading can go faster because readers control how fast they read: there’s no “speaker” to wait for.
- The catch, however, is that the visual system imposes hard limits on how fast we can read. Properties of the eye limit how much can be seen at a time, creating a major bottleneck. The perceptual span—the amount of information that registers during an eye fixation—is surprisingly limited, with only 2 or 3 words clearly visible, at best. Our eyes don’t allow us to take in entire lines of text at a time.
- Reading consists of a succession of fixations (pauses) and saccades (jumps to the next fixation). Most words in texts are fixated at least once, with the exception of short words like *of* and *an*. Many words are skipped when we skim a text, which results in shallower comprehension.
- Good readers average about 4-5 words per second (240-300 words per minute). People do not read faster by making fewer fixations or larger saccades. Rather, faster readers spend less time on each fixation because they recognize and comprehend words more rapidly.
- Reading speed also depends on the difficulty of the text, the reader’s familiarity with the topic, and how deeply the text is read (one’s goal in reading it).

- People have a deep desire to read more quickly, which has led to decades of “speed reading” schemes. These schemes do not work because they are based on false assumptions about skilled reading. Recommendations such as “eliminate the voice in your head” while you read or avoid regressive eye movements (looking back) make reading more difficult, not easier.
- *Anyone* can “read” thousands of words per minute if you don’t have to comprehend very much: just skim. People who claim to read thousands of words per minute *with good comprehension* are either bluffing or deluded. Often, they are trying to profit from their supposedly extraordinary skill. They avoid rigorous comprehension tests, however.
- Changing the way text is presented, as in the RSVP (Rapid Serial Visual Presentation) method, doesn’t improve reading speed, either. The method involves presenting a text word by word in the same location on a screen. This eliminates the need to make eye-movements; however, it also eliminates *advantages* of text, such as the option to vary reading speed, re-read, or jump ahead.
- The amount and variety of reading we do has the biggest impact on reading speed and skill. People who read more are better at recognizing words and understanding texts. They also benefit from more extensive “background knowledge” from having read so much.

### tl;dr

Properties of the visual system limit how fast we can read with good comprehension. The amount of visual information that registers on each fixation is limited by properties of the retina. This limitation can’t be overcome by training yourself to scan the page differently or other “speed reading” tricks. Reading speed and skill depend on the amount and variety of reading. People who read more get better at recognizing words rapidly, but also learn more from the material they read. Having more extensive background knowledge (about a text’s topic, author, genre, and other properties) increases reading skill, which in turn encourages people to read more often. This feedback loop is the mechanism behind skilled reading.

## CHAPTER SECTIONS

1. Introduction – p. 59
2. Eyes Like Yours – p. 62
3. The Perceptual Span – p. 64
4. We Like to Look – p. 68
5. Speed Reading and Reading Speed – p. 70

6. Change the Reader – p. 73

7. Change the Text – p. 79

## **Introduction**

People can read more rapidly than they can comprehend speech. Understanding spoken language depends on what a speaker says and how they say it. Listeners are at the mercy of speakers who decide how rapidly to speak and make other choices that affect how easily they are comprehended. Reading avoids these problems because the full text has already been composed. You, the reader, determine how quickly you read, and you can re-read a section or jump ahead if that helps. People can read popular books at 300 words per minute without strain, but the audiobook versions are recorded at about half that rate. Text also allows the creation and comprehension of more complex language than speech, and so is more than just a minor add-on.

Text offers a lot of advantages, but also comes with limitations that primarily arise from characteristics of the visual system. Our eyes weren't designed for reading. Reading works within the limitations imposed by vision.

## **Eyes Like Yours**

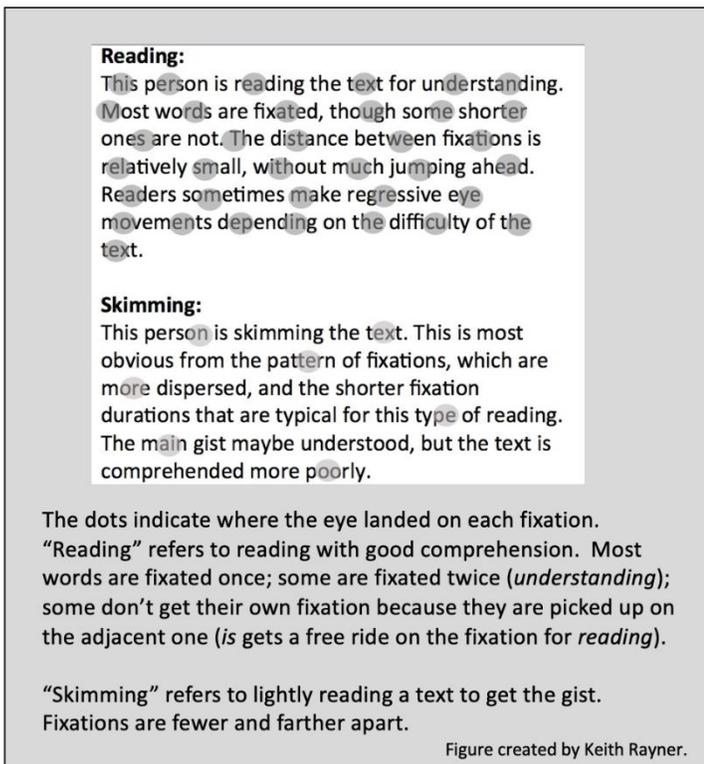
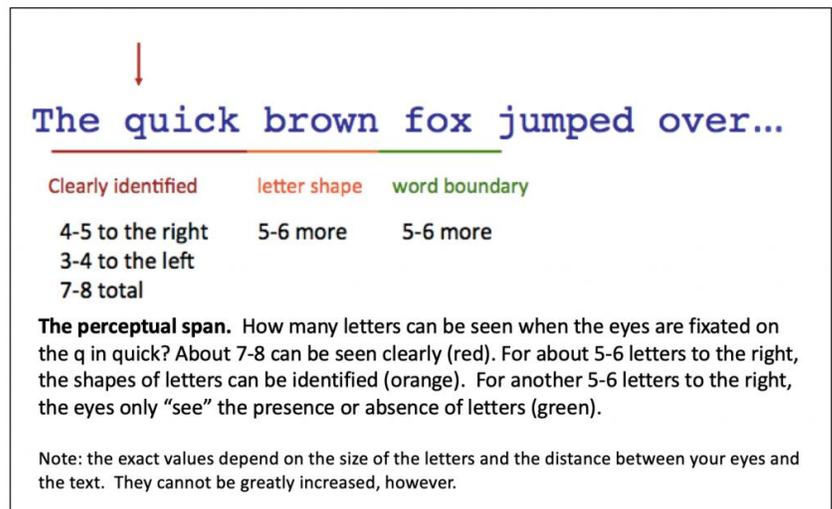
At the “front end,” reading involves processing visual images on the page, recognizing them as symbols, and connecting these symbols with language. The visual part of reading consists of a series of fixations (brief stops) and saccades (jumps to the next fixation). These properties can't be changed: for example, we cannot move our eyes in a smooth, continuous manner across a line of text. That's a prime example of how reading has to work within the strictures of the visual system.

## **The Perceptual Span**

How many letters can be “seen” while fixating at a particular location in a line of text? “Seen” in this case only means that at least some information about the letter is processed by the visual system. The perceptual span is about 15-16 letters to the right of the fixation and about 3-4 letters to the left. In writing systems that are read from right to left, the pattern is reversed.

The letters within this window are not “seen” equally clearly. The ones closest to the fixation point are seen with high acuity. Only partial information is detected for letters that are farther away (see figure). This partial information is used in planning where to fixate next.

Properties of the perceptual span—the fact that the window is limited to about 20 letters; the small number that are seen clearly; the drop-off in acuity further away from fixation—are determined by the distribution of receptor cells on the retina. We can't change this fact about our biology because we can't grow additional cells on our retinas. Nor can corrective lenses improve acuity in the periphery (the orange and green zones in the figure).



## We Like to Look

Fixation durations depend on the difficulty of the word, the difficulty of the text, how closely the text is being read, and the skill of the reader. Content words such as nouns and verbs are fixated longer than function words (pronouns, conjunctions, auxiliary verbs, etc.). Words that we don't use very often (like auxiliary) are fixated longer than very familiar words. Most words are fixated once, some twice. Some short ones get a free ride on another word's fixation.

## Speed Reading and Reading Speed

Many people want to be able to “speed read,” but the truth is that reading speed cannot be greatly increased without compromising comprehension. “Speed reading” techniques only teach people to skim. Researchers have known this for a long time, but the market for “speed reading” products (also marketed under names like mega-reading) persists. In an attempt to drive a stake through these vampire products, I'll explain why speed reading is impossible. These products use two approaches: either change the reader's behavior or change how the text is displayed.

## Change the Reader

Speed reading advocates assert that the secret to reading megafast is changing your “bad” reading habits. Just learn to take in more information at a time; eliminate subvocalization; and stop making regressive eye movements. The first is impossible; the second and third make reading harder.

“Take in more information at a time” defies properties of the visual system described above. You can’t just talk yourself into seeing more letters. The retina isn’t listening!

Subvocalization refers to repeating something to yourself, as when you try to remember a phone number or street address. If we say words to ourselves as we read, it would definitely slow us down, because talking is slower than reading. Speed reading programs therefore include exercises to suppress the “subvocalization” habit. This is a mistake because skilled readers do not “subvocalize”. We do something else: we “hear” words in our minds without having to say them. It is like being able to visualize of what an owl looks like without seeing one. We can also mentally “hear” the owl’s hoot and how it differs from a canary’s song. We can also “hear” the sound of the word owl. “Hearing” the sounds of words in your head (we call this computing a word’s phonological code) is different from saying words to yourself (subvocalizing). Using phonology doesn’t slow reading because there’s no overt speech involved.

We use phonology in silent reading because we can’t help it. With extensive practice the spellings and sounds of words become closely intertwined, down to the neural level. You read a word, you automatically activate information about its phonology and meaning. Poor readers have more difficulty doing making close connections between these codes.

Telling people to avoid regressive eye movements sounds like it should save time, but no. People look back for a reason: because they might not have understood something the first time. Eliminating this option (as with RSVP) makes reading more difficult—unpleasant, even. But speed-reading programs want you to practice this to improve reading. They’ve also got the cause-and-effect backwards. Although everyone makes regressive eye movements, good readers make fewer of them overall because they don’t have to look back as often. You don’t become a better reader by eliminating regressive eye movements; you reduce the number of regressions by becoming a good reader!

## Change the Text

The only other option is to present texts in a format that would allow us to take in more words at a time. With RSVP a text is read without moving the eyes at all. Sounds good, right? Words are

presented (usually one at a time, sometimes more) at a single spot on a screen. The reader has to keep looking at that spot as the words flash by. That means losing control over the rate of transmission—slowing down for harder words, speeding up for easier ones, for example—and the option to re-read. Studied in a research laboratory under optimal conditions, college students can understand texts presented at very high rates (700-800 words per minute), which is impressive. However, peak performance can only be sustained for brief bursts because it requires such a huge amount of effort.

Bottom line: There's no free lunch when it comes to reading speed. Eye-movements aren't the cause of skilled reading or reading impairments; they are determined by how well one reads. Trying to become a skilled reader by changing how you move your eyes is a waste of time. So is using "vision training" to help struggling readers: with rare exceptions, their reading difficulties are not in their eyes. The real recipe for becoming a skilled reader is (a) read a lot; (b) read texts that vary in content, complexity, and style; (c) know a lot about the topic you're reading about!

## STUDIES MENTIONED

Page 74: [\*Bélanger et al.\*](#) (2012)

Page 75: [\*Green & Bavelier\*](#) (2003)

## KEY TERMS

**Fixation** (p. 63) – brief pauses when reading.

**Saccades** (p. 63) – rapid jumps between fixations.

**Regressive eye movements** (p. 63) – occasional backwards saccades, allowing one to reread part of the text.

**Smooth pursuit** (p. 63) – eye movements that allow the eyes move smoothly, without saccades. Can be done if the eyes are tracking a moving object; can't be done while reading.

**Perceptual span** (p. 64) – the "window" around an eye fixation in which letters are processed by the visual system. Only a small number are seen clearly.

**Retina** (p. 66) – the part of the eye that perceives light and converts the light into neural signals that are sent to the brain.

**Fovea** (p. 66) – the part of the retina with the highest visual acuity.

**Cones** (p. 66) – cells in the retina that respond to light, responsible for fine detail and color. Concentrated in the fovea.

**Periphery** (p. 66) – in vision, parts of the retina far from the fovea. Fewer cones, lower acuity.

**Acuity** (p. 66) – sharpness of vision.

**Parafoveal magnification** (p. 67) – a research method in which text in the periphery is made larger and therefore easier to read. Done as an attempt to see if the perceptual span can be increased. The answer: No.

**Subvocalization** (p. 75) – covert speech, saying words aloud very quietly.

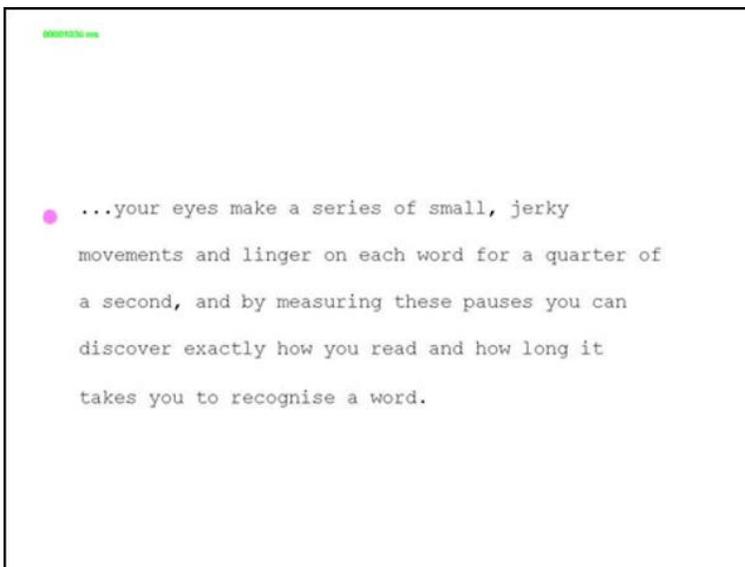
**Subvocalization Fallacy** (p. 75) – the belief that the mental activation of the phonological code is the same as subvocalizing.

**Boustrophedon** (p. 79) – writing from right to left and then from left to right in alternating lines.

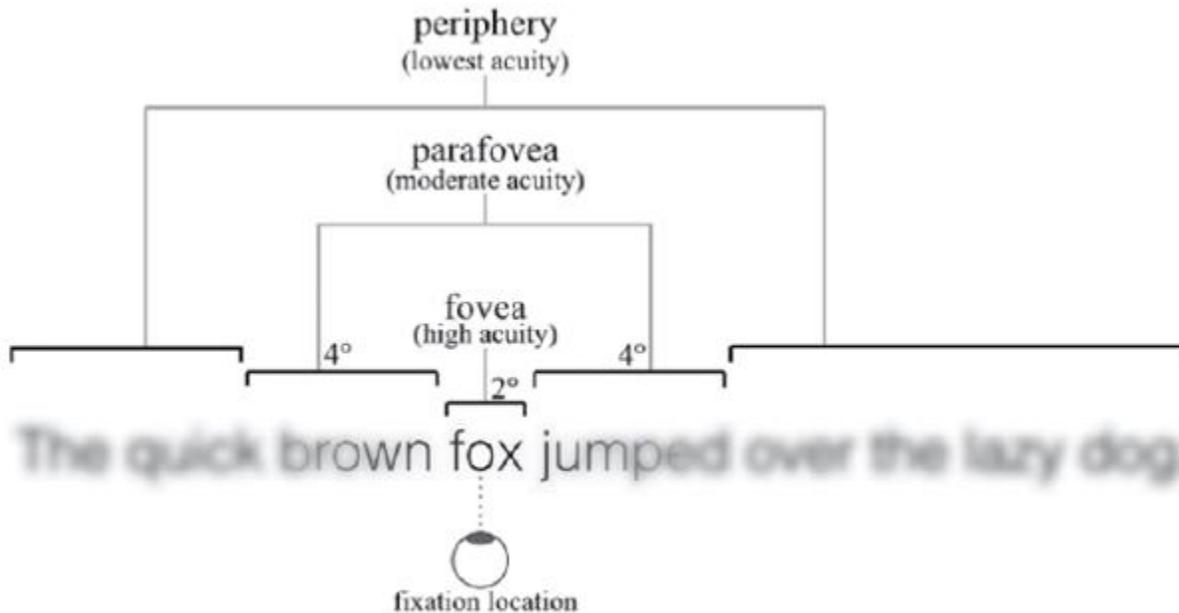
**Rapid Serial Visual Presentation (RSVP)** (p.79) – presenting a written text in a single location, showing single words or combinations of words in attempt to offset the constraints of the visual system.

## RELATED CONTENT

Page 63: [Eye tracking](#)



Page 66: [Eye movements in reading](#)



**Fig. 3.** Diagram illustrating visual acuity across the three regions of the visual field. Acuity decreases continuously as a function of distance from fixation location. The visual field consists of the fovea (center, with highest acuity), parafovea (middle region, with moderate acuity), and periphery (farthest region from fixation, with lowest acuity).

Page 77: ["Eye Tracking Reading Study"](#)

Page 79: [RSVP](#) for The Raven

Page 79: [Try RSVP](#)

## DISCUSSION QUESTIONS

1. The visual system seems to create a bottleneck: the perceptual span is pretty small; we have to make a lot of fixations and saccades. Why would the visual system have evolved in this way? Are these properties more adaptive (helpful) in other kinds of visual situations? [p. 63]
2. OK, so speed reading programs don't work as advertised because they are based on false assumptions. Other products that don't work get taken off the market when customers catch on. Why do you think these programs have survived for so long? [p. 79]

3. After completing the reading speed test are you surprised by your results? Faster or slower than expected? What would have made the text easier to read? (hint: background knowledge) [p. 84]
4. Have we challenged any of your assumptions about what happens when we read? If you are a teacher or parent, have you ever encouraged people to do things like suppress subvocalization or avoid looking back?

## ACTIVITY

Try reading a text using spreeder (an RSVP tool). Vary how fast the text is presented. What's a comfortable presentation rate for you? For how long? It gets tedious doesn't it? You can also present more words at a time (2 or 3, say). Does that help? Would you recommend this method to others?

## HANDY QUOTES

"Reading is the prime example of a technological add-on that extended our capacities beyond their natural limits."

Reading using RSVP "is like having a staring contest with a book."

## 5.

# F U CN RD THS, U CN GT A GD JB N RDNG RSCH

---

## OVERVIEW

- This chapter introduces the idea that language exhibits statistical patterns that are crucial to our ability to learn and use it. "Statistical" means that patterns differ in frequency (how often they occur) and probability (given one pattern, how likely is it to occur with another one?). "Patterns" occur across units such as letters, phonemes, and other building-blocks of written and spoken language. We've learned a lot about these properties of language by analyzing very large samples of texts and speech ("data-mining").
- Languages consist of a hierarchy of different types of elements: letters, phonemes, syllables, morphemes, words, phrases, and more. At each level, a smaller number of elements combines to form a much larger number of units at the next higher level. The 26 letters of the English alphabet, for example, are combined to form many thousands of words. Words, in turn, can be combined to form a nearly infinite number of sentences.
- Language statistics concern the frequencies and co-occurrences of the elements within each level (e.g., how letters combine to spell words) and between levels (e.g., correspondences between the spellings of words and their pronunciations, meanings, morphological and syllabic structure, and other properties).
- Statistical patterns arise from restrictions ("constraints") on how the elements within each level combine. For example, only a very small subset of the millions of possible ways that letters can be combined form actual words. *Talk* is a word, *tilk* could be a word, but *itlk* cannot be a word. Readers know in advance that some combinations of letters will *not* occur. That makes it far easier to recognize the ones that *do*.
- Constraints on how letters combine to spell words mainly arise from the fact that spelling represents the sounds of words. Spoken words are constructed out of units—phonemes—that only occur in certain combinations. Again, many combinations are possible but very few are used. For example, "bl" occurs at the beginning of words but not at the end. Those patterns—and many many others—also occur in the corresponding spellings.

- Statistical patterns also arise from the fact that the various levels of linguistic structure are not independent. The spelling of a word represents more than just its pronunciation. We know that *dress* is the spelling of a word (and that *dsers* is not), but also that *dress* is a syllable with a particular pronunciation, similar to *press* and *dross*; and a morpheme that occurs in related words such as *dresses* and *undress*. Spelling also provides cues to units such as syllables. For example, if a word contains *nv*, there must be a syllable boundary between the letters, as in *canvas*. Language is full of such patterns, at every level.
- Statistical patterns arise from many other sources, too. Some sounds are used more often than others because they are easier to pronounce; how often words are used depends on what we choose to talk about—chairs, for example, more often than ottomans—which in turn reflects patterns that exist in the world. An author's distinctive style is reflected in statistical properties of the texts they write—which words are used and how they are combined.
- We learn about these statistical patterns through our use of spoken and written language. Every time we read a text or hear a sentence, we update our language statistics. Statistical learning refers to people's capacity to learn complex statistical patterns, as in language and reading, but also face and object recognition, making judgments about people's beliefs and intentions, and many other everyday phenomena.
- Statistical learning is implicit: it occurs in the background, without conscious awareness, as we engage in activities such as reading and listening. Implicit learning contrasts with explicit learning, as occurs when learning facts such as the names of the 50 states or the 6 types of syllables in English.
- The scrambled text known as the Cmabrigde reading hoax illustrates that people can use their knowledge of statistical patterns to understand texts with many letters deleted. So does the title of this chapter. Notice, however, that reading these texts is harder and slower than normal reading. Skilled, fluent reading depends on knowledge of the actual spellings of words, and other language statistics.
- The mechanism involved in combining different kinds of statistical patterns is called "constraint satisfaction." It's an important component of human intelligence.

## tl;dr

This chapter may seem a little obscure to some people--sorry! It concerns the basic structure of written and spoken language, and how that makes reading and speech possible. There are a couple of big ideas to hold onto. The spellings of words consist of combinations of letters that exhibit statistical structure: some don't occur at all, some only occur in certain positions, some occur much more often than others, etc. The same is true of how phonemes combine to form spoken words, and words combine to form sentences—it's true about every level of language structure. Patterns also exist between levels: letters and phonemes, for example.

That is because the components of language are related to each other, not independent. People are exceptionally good statistical learners—we do it all the time. Most of this learning occurs unconsciously, as we read or speak (though explicit instruction is required to get the system rolling). These observations have important implications for learning to read and the role of instruction. Finally, the brain is very good at combining these statistics, the "constraint satisfaction" concept discussed at the end. Bits of information that aren't very informative in isolation suddenly become very helpful when combined. See the examples for clear illustrations (including one from "Wheel of Fortune").

## CHAPTER SECTIONS

1. Introduction - p. 85
2. Department of Redundancy Department - p. 88
3. Varieties of Orthographic Expertise - p. 92
4. Why Statistical Knowledge Is Powerful - p. 96

### Introduction

Data-mining ("Big Data") is the process of detecting patterns that are hidden in masses of data. Modern Artificial Intelligence systems discover such patterns in learning to perform tasks such as speech, face, and handwriting recognition; predicting people's movie and clothing preferences; and many others.

Language exhibits statistical patterns at every level. Learning and using a language is a "big data" problem for humans. We collect data about the patterns in language implicitly, in the background, as we use written and spoken language. Becoming a reader involves learning about the statistical properties of print—the frequencies of letters and how they combine—and about the mappings between print (orthography) and the other major linguistic codes: sound (phonology) and meaning (semantics).

## Department of Redundancy Department

Written language is redundant: only some patterns occur in words, and the ways they combine are highly constrained. This knowledge makes it easier to recognize words. Captcha security systems make use of this. If you can recognize some of the distorted letters in a captcha, you can fill in others because only some combinations form legal words.

## Varieties of Orthographic Expertise

We read in order to understand texts, but in doing so we gain information about spelling and other aspects of language. Instruction and practice are important for beginning readers, but most of our knowledge of print isn't gained through instruction or study. There is an interesting exception, however: competitive Scrabble players. Scrabble requires knowing which combinations of letters form legal words. Elite players study lists of spelling patterns in order to gain this knowledge. It is a peculiar skill because, unlike readers, they do not need to know the meanings of the words. Most of us learn about spelling through the act of reading. Every time we read, we update this knowledge. To get this type of learning underway, beginning readers need instruction, practice, and feedback. The relatively small number of patterns that can be explicitly taught provide the foundation for gaining additional expertise through reading and understanding texts.

Readers know a vast amount about the spelling system, including statistical patterns that we use without conscious awareness. The power of this knowledge is demonstrated by an Internet meme, the Cmabrigde Hoax. The text could be read even though letters had been deleted. We can fill in the missing information. However, the further claim that this shows that readers don't pay close attention to the letters in words or their order is wrong. The scrambled text is harder to read than a normal text, and the scrambling had to be done very carefully and selectively to make it readable at all.

## Why Statistical Knowledge Is Powerful

Language statistics tell us what is likely to occur. TAL is likely to be followed by K or L, and maybe by C or E, but not by M or P, for example. This kind of knowledge is powerful because it holds for every type of linguistic knowledge, and because the levels are related, not independent. Vocabulary—knowledge of words—includes information about the ways that letters and phonemes combine; knowing about orthographic and phonological patterns helps in learning new words. What is learned about one thing carries information about the others. That's how people manage to learn all this stuff even though it can't all be taught in class. The final bit is that our brains are very good at combining these bits of information. That is the "constraint

satisfaction" idea discussed at the end of the chapter, and it is important. Bits of information aren't very helpful in isolation, but become very important when they are combined.

## STUDIES MENTIONED

- Page 85: [Markowitz & Hancock](#) (2014)  
Page 87: [Romberg & Saffran](#) (2010)  
Page 92: [Blough](#) (1982)  
Page 92: [Grainger et al.](#) (2012)  
Page 93: [Tuffiash, Roring, & Ericsson](#) (2007)  
Page 93: [Hargreaves et al.](#) (2012)  
Page 94: [Maguire et al.](#) (2000)  
Page 96: [Ashok et al.](#) (2013)  
Page 97: [Seidenberg & MacDonald](#) (1999)  
Page 97: [McClelland & Rumelhart](#) (1981)

## KEY TERMS

**Stylometry** – the statistical analysis of variations in literary style between writers or genres  
- p. 85

**Implicit learning** – the acquisition of knowledge in an incidental manner, without conscious awareness - p. 87

**Explicit learning** – the acquisition of knowledge in an intentional manner, with conscious awareness - p. 87

**Pseudowords** – a word in speech or text that obeys the rules of a particular language and appears to be an actual word but contains no semantics - p. 92

**Lexical decision** – a task that measures how quickly participants are able to recognize stimuli as words or nonwords - p. 92

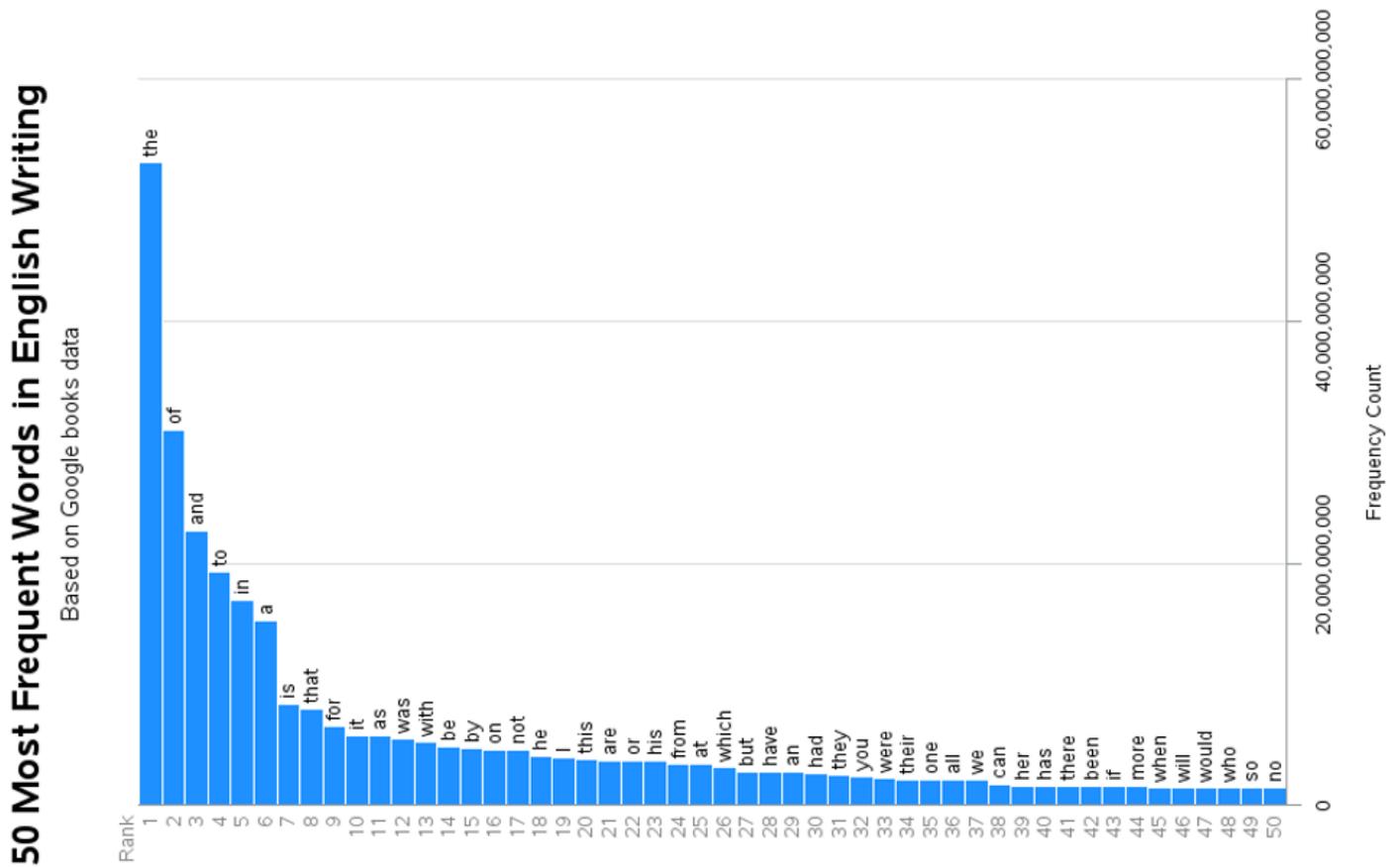
**Anagrams** – all of the words that can be formed from a specific group of letters - p. 93

**Bigram transposition** – reversing the order of two letters, a type of transposition that is easy to override - p. 94

**Constraint satisfaction** – removing the choices that are not possible so that the best ones remain, also known as process of elimination - p. 97

## RELATED CONTENT

Word frequency long tail distribution - p. 89



[Peter Norvig's orthographic statistics](#) - p. 91

Bigram transposition - p. 94



[Wheel of Fortune constraint satisfaction](#) - p. 97

## DISCUSSION QUESTIONS

1. If you had seen the Cmabrigde hoax before, what was your intuition about why you could read the passage? Has your understanding changed after reading this chapter? [p. 88]
2. How could knowledge of orthographic statistics help children who are learning to read and spell? [p. 89]
3. How do statistical regularities help skilled readers read? [p. 96]
4. Before reading the next chapter, come up with a list of all the skills an advanced reader must have.

## ACTIVITY

Create two lists of made-up words – one that contains pseudowords (words that could exist in English) and one that contains words that could not occur in English. How similar are the words in the first list to real English words? How many letters would need to be changed to make a real word? What patterns did you violate to create the second list? Was either list easier to create?

## HANDY QUOTES

“Learning to read is the process of acquiring the several types of statistical knowledge that support rapid and efficient comprehension, starting with phonological structure, orthographic structure, the mappings between orthography and phonology, vocabulary, and grammar.”

“The fact that most letter combinations do not occur make it easier to recognize the ones that do...The spelling patterns that *are* used are a highly nonrandom sample of millions of possibilities.”

## 6.

# BECOMING A READER

---

## OVERVIEW

- This chapter is about the precursors of reading: the experiences of very young children—experiences that later affect learning to read. It focuses on spoken language experience and children’s initial experiences with print (letters, letter names, sounds associated with letters).
- Children always learn a spoken (or signed) language before they learn to read. Speech is the initial source of knowledge about language and things in the world we use language to communicate about. Reading lags behind speech but then overtakes it. For people who are literate, reading becomes the vehicle for expanding knowledge of language and the world. Texts contain words, expressions, and content that do not occur in everyday speech.
- Learning to talk involves learning spoken words and their meanings. As a child’s vocabulary expands, they begin to discover that spoken words have parts. For example, from rhymes like “book” and “look” children learn about the different initial sounds (the onsets “b” and “l”), which are phonemes, and the “ook” sound (a unit called the rime), which occurs in many words. Discovering the components of spoken words (“phonological awareness”) is important for reading.
- Why do letters have names like “bee” and “aitch”? The names are handy for referring to letters, but they do something more. Letters such as B can be written many ways. The letter name helps the child treat all of the variations as members of the same category, the letter B.
- Vocabulary is more than just the number of words a child knows. The first question to ask is whether a child knows a word at all, e.g., that “shoe” refers to a certain kind of object. A further question is how much the child knows about “shoe”: kinds of shoes, horse shoes, words that are commonly used with shoe (like box), where shoes are found, what they are made of, the fact that shoe can also be a verb, etc. Eventually they learn that shoe tends to occur in certain kinds of sentences and expressions. Because of differences in children’s experience, vocabulary size differs as does “lexical quality”—how much information they associate with the word.

- Vocabulary acquisition raises a deep question: children know more than a thousand words by the time they start school. Only a small percentage of these words were explicitly taught. How do children learn so many words so quickly without being taught each one? The answer: language statistics. Words overlap in meaning; learning about one prepares the way for learning about related words. Words that are similar in meaning also occur in similar positions in words. These statistical patterns provide strong clues to the meanings of new words, allowing rapid acquisition with a minimum of explicit guidance.
- Reading to children is beneficial for many reasons and it is important to encourage. However, it usually isn't the same as teaching children to read. Shared reading mostly focuses on the meaning of the text and the shared experience, not instruction about spelling or grammar (there are exceptions, of course). Though it is very important, reading to children does not inoculate them from a condition such as dyslexia. Nor does it guarantee that they will become good readers: instruction also matters!
- Reading books aloud to children has another function that isn't usually recognized: children's books include words and expressions that differ from everyday speech. Reading with young children provides a way to expand their knowledge of language and the world.
- Spoken language development and early exposure to print bring the child to the brink of reading. When do children start to read? The exact timetable varies depending on the child and the environment. Some children get off to a faster start than others for a variety of reasons (e.g., they were taught to read at home before starting school). It's more important to consider how far children have progressed through the first several grades.
- The "Simple View of Reading" is a useful description of the situation at the onset of reading. The child already knows about language (and the world) from speech. The new part that must be learned for reading involves learning about print, specifically how words that are known from speech are represented in this visual-graphical code.
- Being able to pronounce words aloud is an important step in early reading. If a word is pronounced correctly, a child can recognize it as a word they know from speech. They can then comprehend it just like when it is spoken to them. Eventually, reading aloud transitions into generating the sound of a word mentally, "in the mind's ear." This is much faster than pronouncing the word aloud. This internal representation of the pronunciation/sound of a word is called its phonological code. The use of phonology to read words is not limited to beginning readers. Skilled readers use mental phonology as well.

- Learning the correspondences between spelling and pronunciation is important because they allow words to be recognized more rapidly than other methods such as trying to directly associate the spellings of words with their meanings, or guessing words from the context in which they occur.
- The spelling of a word is a sequence of discrete letters. Learning the correspondences between spelling and sound is easier if the reader treats spoken words as consisting of a discrete sequence of sounds—phonemes. Spoken words don't *really* consist of discrete sounds like beads on a string, but it is useful to treat them as if they do. This abstraction then allows units in the written code to correspond to units in the spoken code. That makes the correspondences easier to learn.
- The idea that letters (or digraphs like SH) represent phonemes is sometimes called the alphabetic principle. This “principle” summarizes something children learn gradually. It reflects how spelling and sound are represented in the brain—each one having shaped the other during the course of learning.
- “Phonemic awareness” refers to the ability to treat words as though they consist of discrete sounds. The term “awareness” is confusing. Children do not need to be consciously aware of what a phoneme is. They do not need to learn the term at all. The instructional goal is to lead them into treating words phonemically, as demonstrated by performance on tasks such as deciding if two words begin with the same sound, end with the same sound, counting the number of sounds in a word, etc.
- These first steps in learning to read are challenging for many children. Instruction therefore matters. All writing systems have arbitrary foundations: the types of symbols that are used (e.g., letters vs. characters); the graphical shapes of the letters; which letters happen to be included in an alphabet; which sounds the letters represent. Learning this is complicated by the cross-modal nature of the associations between spelling and sound, and the complexity of the mappings in English.

## tl;dr

Children’s early experience with spoken language, and their knowledge of the things we use language to communicate about, have enormous impact on reading. Reading depends on speech. Children’s early language experiences vary in ways that can help or hinder later reading. Reading to children exposes them to a broader range of words and expressions and expands their knowledge of the world beyond their immediate experience. With few exceptions, reading requires additional instruction, mainly because of the characteristics of the printed code and the nature of the correspondences between print and speech.

## CHAPTER SECTIONS

1. Introduction – p. 101
2. Proposed Requirements for Licensure as a Certified Skilled Reader – p. 102
3. Reading Develops – p. 104
4. Why Do Letters Have Names? – p. 107
5. Vocabulary: It's Not Just Number of Words – p. 109
6. Reading to Children: Necessary but Not Sufficient – p. 113
7. Finishing the /h/-/a/-/t/ – p. 117

### Introduction

Children's movement towards literacy is dependent on their experiences at home and at school, which are in turn, reliant on the adults in these settings. For decades, how children should be taught to read has not been a trivial issue. The disputes surrounding this question gave rise to the reading wars.

### Proposed Requirements for Licensure as a Certified Skilled Reader

A skilled reader must be able to recognize and understand a large variety of words, sentences, and sentence structures. Skilled readers must also be able to infer meanings of unfamiliar words and structures, as well as possess the ability to analyze texts on a deeper level.

### Reading Develops

Because reading skill must develop, the way to become a skilled reader is not by imitating one. The reading skill of a particular individual depends on where they are on the reading trajectory. Advanced skills for a particular age group may be only proficient for another group. The challenge facing new readers includes figuring out how print represents speech and how to comprehend language in a new form. As children learn a language, they first perceive words as auditory blobs. Rhyming helps children learn to isolate specific phonemes. During this time, the process of lexical restructuring occurs. This theory asserts that during children's rapid growth in vocabulary, they learn to isolate phonemes and that these phonemes are what make up words.

### Why Do Letters Have Names?

Names of letters allow for words to be spelled aloud and in many cases provide hints about pronunciation. In some cases, the names and sounds are related, although the relationship is not perfect. Letter names also give children labels for each of the 26 categories they must learn. Having a name for a category facilitates learning and helps children classify a letter they come correctly, regardless of stylistic form.

### **Vocabulary: It's Not Just Number of Words**

Vocabulary can vary in its lexical quality – the amount of information that is associated with a particular word beyond the basics of orthography, phonology, and meaning. Words are tied to expectations based on our experiences with both language and the world in general. How children develop large vocabularies quickly is similar to how children learn language, more generally: through big data. One can infer the meaning of a new word because words that are similar in meaning tend to occur in the same contexts. Explicit instruction can accelerate vocabulary acquisition by updating the statistical patterns that are relevant to many other words.

### **Reading to Children: Necessary but Not Sufficient**

Reading to children offers benefits outside of the realm of language. Children can learn to enjoy the experience of reading, learn about places and things otherwise not in their daily life, see new pictures, and hear exciting stories. Print exposure leads to larger, richer vocabularies and knowledge of syntactic structure, which in turn promotes better reading, leading to further development of said knowledge. However, most children will not learn to read through observing another person reading. Socioeconomic disparities influence the frequency of book reading at home. Low-income families are less able to buy books and live in neighborhoods with fewer public libraries which contain books of lower quality. The caregiver's ability to read and their availability also influence the amount of book reading. Several interventions have been implemented to try to alleviate the effects of socioeconomic status, such as Reach Out and Read, a program that has helped distribute books to millions of children.

### **Finishing the /h/-/a/-/t/**

Children must learn to link print to their preexisting knowledge of spoken language. Initially, they must learn that graphemes are used to represent phonemes. This is difficult to learn because (1) learning about phonemes requires exposure to print – children have no preexisting concept of phonemes, (2) the relationship is arbitrary– the appearance of letters is unrelated to pronunciation, and (3) the associations are cross-modal – they rely on both the visual and auditory modalities. Over time, children gain phonemic awareness, the understanding that words consist of phonemes. Children may not be conscious of this understanding or be able to articulate it, but they can demonstrate it in tasks.

## STUDIES MENTIONED

- Page 105: [\*Beck & Juel\*](#) (1995)  
Page 106: [\*Browman & Goldstein\*](#) (1990)  
Page 106: [\*Werker & Tees\*](#) (1999)  
Page 107: [\*Metsala & Walley\*](#) (1998)  
Page 108: [\*McCardle, Scarborough, & Catts\*](#) (2001)  
Page 109: [\*Lupyan, Rakison, & McClelland\*](#) (2007)  
Page 110: [\*Perfetti\*](#) (2007)  
Page 111: [\*MacDonald, Pearlmutter, & Seidenberg\*](#) (1994)  
Page 111: [\*Seidenberg\*](#) (1997)  
Page 111: [\*Beck & McKeown\*](#) (1991)  
Page 111: [\*Hoff\*](#) (2003)

## KEY TERMS

**Reading wars** – the debate surrounding the best way to teaching reading – p. 102

**Rime** – the part of a syllable which consists of its vowel and any consonant sounds that follow – p. 107

**Lexical restructuring** – the impact learning to read has on the specific lexical patterns – p. 107

**Exemplar** – the typical or ideal example for a particular for a certain concept – p. 108

**Lexical quality** – the quality of representations affects reading skill and comprehension. High lexical quality allows for fast and accurate retrieval, whereas low quality elicits specific word-related difficulties – p. 110

**Bootstrapping** – in this context, the idea that children acquire vocabulary by using semantic context to learn the meaning of new words – p. 113

**Fast mapping** – the process by which a child learns a new word very quickly, often after only one exposure – p. 113

**“a simple view of reading”** – early reading is comprised of two components – print knowledge and comprehension. Beginning readers must learn to link written language to their existing knowledge about spoken language – p. 118

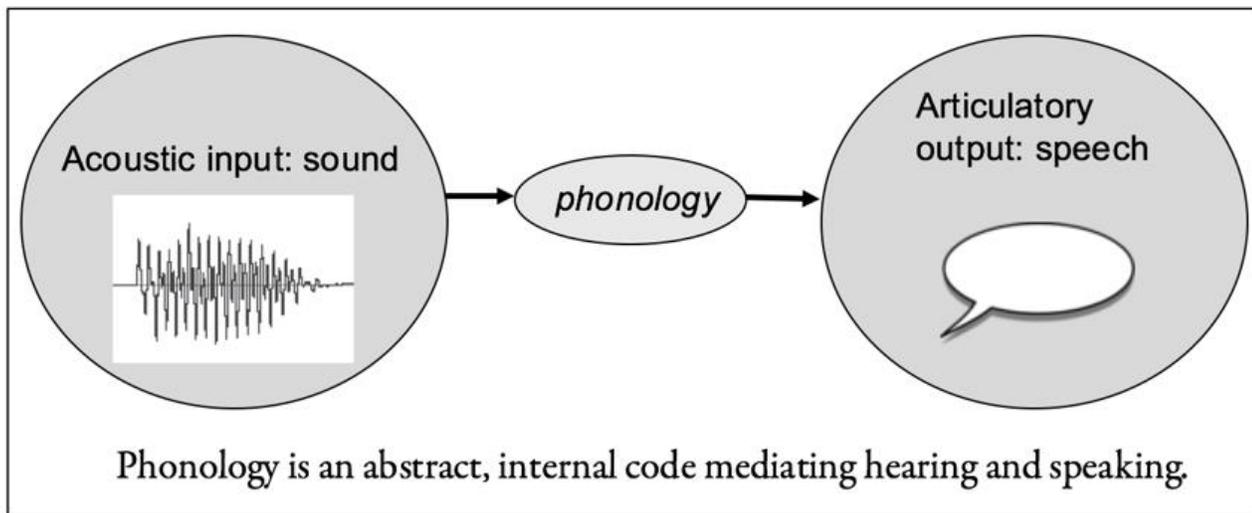
**Alphabetic principle** – understanding that spoken words are decomposed into phonemes, and that the letters in written words represent the phonemes in spoken words when spoken words are represented in text – p. 119

**Cross-modal** – perception across two or more different modalities, e.g. vision and audition – p. 120

**Phonemic awareness** – the ability to recognize words as containing individual units of sound, or phonemes – p. 121

## RELATED CONTENT

The two sides of the linguistic coin: Hearing → Phonology → Speaking – p. 106



[Reach Out and Read](#) – p. 116

[“Pre-readers can use bits they have learned about letters to correctly guess some words ...”](#) – p. 118

## DISCUSSION QUESTIONS

1. Compare the list you created at the end of the last chapter to the list on p. 102. Discuss the overlaps and differences between the two lists.
2. What are the benefits of knowing letter names for children learning to read? [p. 108]
3. How might lexical quality be affected by socioeconomic factors? How could these potential differences be reduced? [p. 110]
4. Discuss the benefits of reading to children. How might socioeconomic factors play into a child’s long-term development and reading achievement? [p. 113]

## ACTIVITY

Make a list of activities that could be used to teach young children about the alphabetic principle. Consider that children need to understand that words are made up of letters, that letters represent the sounds of speech, and what specific sounds different graphemes represent. What types of activities on your list would be most effective for teaching each of these different aspects?

## HANDY QUOTES

“Two important consequences follow from the fact that reading follows a developmental trajectory. First, the way to become a skilled reader is not by emulating one...The second consequence is that what it means to be a skilled reader depends on where an individual is on the trajectory.”

“It should now be clear why becoming alphabetic is a major hurdle that requires instruction, feedback, and practice. The child has to think phonemically, which involves both phonology and orthography, and learn arbitrary cross-modal associations between graphemes and phonemes.”