

discovering speech, words, and mind



dani byrd and toben h. mintz

 WILEY-BLACKWELL

www

Discovering Speech, Words, and Mind

Dani Byrd
and
Toben H. Mintz

© website at www.discoveringspeech.wiley.com



 **WILEY-BLACKWELL**

A John Wiley & Sons, Ltd., Publication

This edition first published 2010
© 2010 Dani Byrd and Toben H. Mintz

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical, and Medical business to form Wiley-Blackwell.

Registered Office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex,
PO19 8SQ, United Kingdom

Editorial Offices

350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

The right of Dani Byrd and Toben H. Mintz to be identified as the authors of this work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Byrd, Dani.

Discovering speech, words, and mind / Dani Byrd and Toben H. Mintz.
p. cm.

Includes bibliographical references and index.

ISBN 978-1-4051-5798-8 (hardcover : alk. paper) – ISBN 978-1-4051-5799-5
(pbk. : alk. paper) 1. Oral communication. 2. Speech. 3. Speech perception.
4. Language and languages. I. Mintz, Toben H. II. Title.

P95.B97 2010
302.2'242–dc22

2009039198

A catalogue record for this book is available from the British Library.

Set in 11/13pt Palatino by Graphicraft Limited, Hong Kong

This book is dedicated to
James Byrd, who didn't live to see it,
and
Benjamin and Anna Siegel

Acknowledgments

There are many people who have helped us. All errors are of course our own. We thank Sophie Ambrose, William Badecker, Freddie Bell-Berti, Hagit Borer, Susie Choi, Sandy Disner, Oliver Foellmer, Teruhiko Fukaya, Louis Goldstein, Ernest Greene, Roberto Mayoral Hernandez, Yu-Chi Huang, Aaron Jacobs, Keith Johnson, Elsi Kaiser, Abigail Kaun, David Li, Frank Manis, Michal Martinez, Franc Marušič, Tatjana Marvin, Emi Mukai, Emily Nava, Laura Siegel, Linda Siegel, Reid Swanson, Jean-Roger Vergnaud, Rachel Walker, Hao Wang, Kie Zuraw, and the Bread and Porridge restaurant in Santa Monica. And a special thanks to Barry Schein and Maryellen MacDonald who conceived the course we've so enjoyed teaching.

Contents

Chapter 1 Human Language as a Scientific Phenomenon	1
Section 1: The specialness of language	1
Section 2: The study of language as a cognitive science	11
Chapter 2 Speaking, Sound, and Hearing	23
Section 1: Speaking and transcribing	23
Section 2: The sound of speech	51
Section 3: Hearing	64
Chapter 3 Phonetic Diversity in the World's Languages and Diversity Among Dialects of English	72
Section 1: Sounds of the world's languages	72
Section 2: English in the United States and around the world	89
Chapter 4 The Scientific Method and Experimental Design	98
Chapter 5 Speech Perception	114
Section 1: The lack of invariance and the use of cues in speech perception	114
Section 2: Ways in which speech perception might be special	127
Section 3: Discovering words	150
Chapter 6 Word Recognition	159
Chapter 7 Phonological Units and Phonological Patterning	182

Chapter 8 Word Form and Function	206
Chapter 9 Sign Languages	224
Chapter 10 Language and the Brain	243
Chapter 11 Language, Speech, and Hearing Disorders	266
Chapter 12 Reading and Dyslexia	282
Appendix – IPA Chart	303
Index	304

® Accompanying web material can be found at
www.discoveringspeech.wiley.com

Chapter 1

Human Language as a Scientific Phenomenon

Section 1: The Specialness of Language

What Is Special About Language

Language is our most important and universal communication medium. As humans, we rely on our capacity to communicate with one another and to “speak our mind.” However, we almost never stop to consider that these abilities are made possible by three facts: we can move our bodies in highly skilled ways; these skilled movements create physical changes in the environment that our senses can apprehend; our brains allow the development and use of a complex system of structuring information for expression to other individuals. As scientists, it is fascinating to turn a scientific eye and scientific tools to studying each of these aspects of language. One reason for this fascination is precisely because our language feats are accomplished with ease and no real awareness on our part of their intricate and structured nature. Throughout this book, we will investigate both the nature of language and our human linguistic abilities.

All healthy humans, and only humans, are born capable of using language, and people acquire these abilities simply through exposure without any overt instruction. You may have to take lessons to learn to play the clarinet, but no children need “do this—don’t do that” lessons to learn their first language. Children gain this ability simply through their normal course of development and interaction with other speakers of a language, just as infants learn visual depth perception or toddlers learn to walk.

Language’s capacity for transmitting information is unrivaled. We’ve all heard the expression “a picture is worth a thousand words.” But while humans can transmit information to one another through facial expression or visual constructions or touch or even smell, none of these can transmit

the enormous amount of detail and specificity that is possible with the use of language. Language can also be tremendously evocative – a poem or story or conversation with a boyfriend or parent can move one to tears, hysterical laughter, an intense insight, or a visceral opinion. Not only does language succeed remarkably in transmitting information from one individual to another at a particular time, it permits the continuity of culture over time through oral, written, and now digital recording of knowledge.

For the scientist newly come to the study of language (perhaps like yourself), an objective quantitative and experimental approach to investigating language can prove challenging for exactly the same reasons that make language an interesting object of study: namely, most people – other than your rare linguist or psychologist – simply don't think much about speaking language. Certainly all of us – even language scientists – learn and use language effortlessly. So as *scientists* of language, we will need to become objectively aware of aspects of language that we normally pay no attention to. To dissect its properties as scientists will require us to suspend the preconceptions or biases we may have about how people speak and also require us to be willing to learn about tools scientists use to investigate human behavior – physical measures of the world and human behavior, techniques for experimentation, and critical analysis of data.

The Speech Chain

Let's consider an extremely simple exchange of information through language. Suppose you are at a party (and are over the legal drinking age) and further suppose that the party has run out of beer. It might occur to you that more beer is needed. Seeing as how friends, however close, do not in fact have ESP, in order to express this desire for more beer, you will need to formulate a message and speak it. In addition to knowing the **proposition** or idea that you wish to express, you will now need to select and combine words to express this thought – in English these words might be *we*, *need*, *more*, and *beer*. However, the particular language you speak, in addition to determining the sounds used to form each word (and even how many words are needed) also determines how these words are to be combined. A linguist would call sensible combinations – *we need more beer* – **grammatical** and term nonsensible combinations – *more need beer we* – **ungrammatical**. (This is a specialized use of the term grammatical having nothing to do with the schoolroom rules.) Now, mind you, a thirsty partygoer is nowhere near finished yet. Having selected the words and put them into a sensible combination, our partygoer must convey them to someone, preferably someone having money, at the party. To do this, the partygoer will have to move his or her mouth or **vocal tract** in intricate fashion using the vocal folds, tongue, jaw, nasal port, and lips. Each word

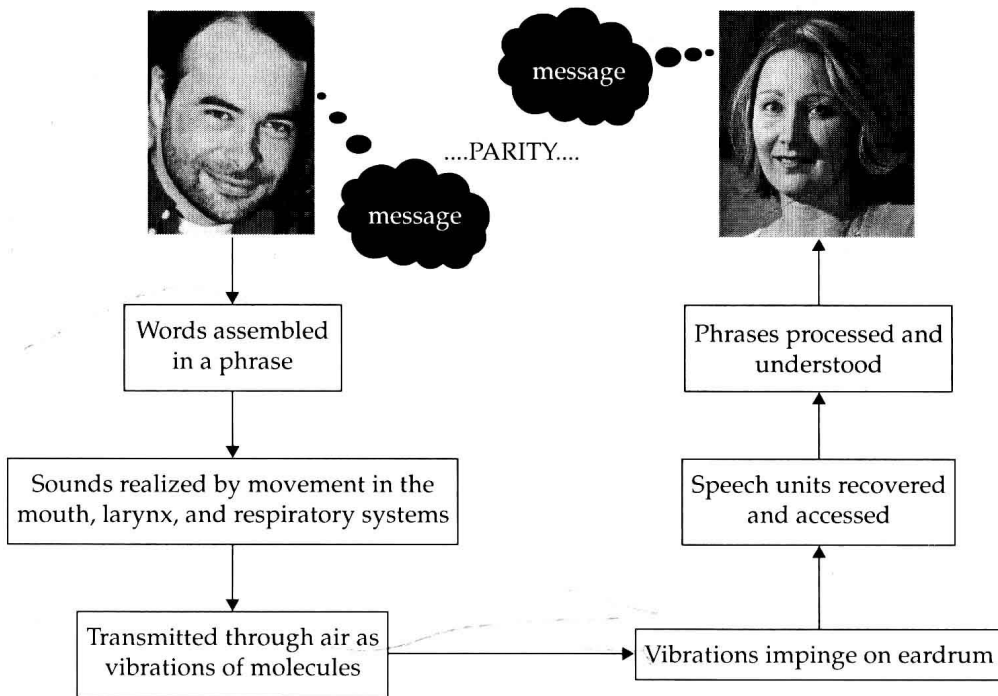


Figure 1.1 The speech chain

is associated with a set of complexly organized movements that cause sound waves to propagate or travel through the air. If the partygoer has been paying attention to where his or her friends are at the party, these sound waves will come into contact with the eardrum of a nearby listener. The listener's brain will respond to the nerve signals created in the inner ear by activating the words in their mind that are consistent with the physical acoustic signal that they just encountered. From these words, the listener will be able to reconstruct the intended message. When the listener recovers the message, and the speaker's intended message matches the one received by the listener, we say that **parity** has been achieved. This is a primary goal in communication.

How can such a complex behavior be so intuitive? The human brain has an inborn or innate capacity for language development and use. This means that it has neural structures whose functionality is, or becomes, specialized for language. Just as humans have evolved to have upright gait, stereoscopic vision, and fine manual manipulation abilities, they have evolved to use language. In addition to the specialization of the human brain, these evolutionary adaptations have also included anatomical and functional characteristics of the vocal tract and auditory system.

It's important to realize that while we have been drawing a picture of *spoken* language, the brain and body can demonstrate the same important

properties with language in another modality, namely *signed* language. Language does not necessitate the use of the vocal tract, sound, and the auditory system; it merely requires that humans act in the environment so as to create a structured, information-carrying signal that other humans can perceive well and with parity. Just like spoken languages, signed language also requires movement of the body in a highly skilled way; these skilled movements create physical changes in the environment that our senses can apprehend (reflection of light off the body sensed by the retina), and our brains allow the development and use of a complex system of structuring information in signed languages. In fact, there is a strong correspondence between the parts of brain that process spoken language through vision (i.e., during reading) and the parts of the brain that process signed language (necessarily via vision).

Language Knowledge

What kinds of scientists study this phenomenon of language? The answer is, fortunately, many different kinds of scientists: linguists, psychologists, physiologists, physicists, physicians, engineers, computer scientists, speech-language pathologists, and educators, for example. In this text, we will use bits of information from a variety of scientific arenas to consider the questions of what knowledge is acquired when a child learns language, how this learning is accomplished, and how speaking and understanding take place.

Linguists describe and/or model what people (often subconsciously) *know* that allows them to speak their language. One piece of knowledge that this includes is the set of sounds used in their language and precisely how to articulate and coordinate them. It's clear that people become highly skilled at making a particular set of sounds or vocal tract actions because when we are confronted with a sound or sound-sequence that's not in our own language – such as the sound at the end of *Bach* or at the end of *Favre* – we often substitute other sounds or combinations – you've heard these pronounced, no doubt, as [bak] or [farv]. (We will be using square brackets in this book to indicate sound pronunciation, known as **transcription**.)

Another type of knowledge that humans have about their language is the connection between a certain sound pattern and its specific word meaning. Like the particular sounds in one's language, the meaning of myriad sound sequences – words – must be learned: the sounds “eegl” means a predatory bird in English and the same sequence of sounds means a hedgehog in German. Humans must learn the mapping between sound and meaning that exists in the particular language(s) they are learning, because a child comes into the world prepared to learn *any* human language. Lastly, a person must know how words may be combined with one

another in ordered structures and what those combinatorial structures mean. So in English, objects follow verbs, but in Navajo, Japanese, Basque, and Hopi, they precede them. In addition to needing to learn all of this – and apparently doing it with amazing grace and speed by the early years of life – a language learner will also need to learn the appropriate patterns of interaction among participants in a conversation or discourse. We use the term **linguistic competence** to refer to all the myriad patterns that a language user knows that allow for the production and comprehension of language. A person's linguistic competence is the idealized body of knowledge of the structures, sequences, and organizations that may or may not occur in his or her language and their relation to meaning. When linguists want to refer to this body of specialized knowledge, they use the term **grammar**. The term grammar does not mean to a linguist what it means to a grade school English teacher. For a linguist, "grammar" is used to mean a theoretical or formal description of linguistic competence or certain aspects of competence.

Languages Change

One preconception we might have about language is that it exists in some pristine form in the minds of a particularly prestigious speaker or group of speakers. In fact, different forms of a language occur in a variety of socio-economic and geographic groups. And every instantiation of language in the mind of any individual is equally valid and worthy of scientific study. Language also changes with each generation of speakers. You don't speak in the same way as your parents and grandparents; nor will your children use the same form of language that you use. Some of these changes found in a particular generation are ephemeral or short-lived. Will the once derogative term "geek" leave the language in the next generation, or perhaps become a positive term for a technologically sophisticated person? Some changes to a language will persist and become incorporated into the more widely spoken language. Believe it or not, the term "email" was once actually limited to academics and technophiles. The fact that language is always changing means that your linguistic competence will be different than that of your parents and children.

You can no doubt think of examples of geographic differences in vocabulary – "pop" versus "soda," "fountain" versus "bubbler." Can you think of pronunciation variants around the United States? In the West, where the authors of this book live, the words *cot* and *caught* are generally pronounced the same; however, in the Northeast, they have different vowels. Likewise *pull* and *pool* and *which* and *witch* are often homophonous in the West, but certainly not elsewhere. Generation-related variation in pronunciation can be harder to pin down, but we find that most undergraduates

pronounce “The End” with the “the” rhyming with “duh” not with “dee.” However, most older folks prefer the other pronunciation for the article when it occurs before a word starting with a vowel. A language change currently in progress relates to the grammaticality (used in the *linguistic* sense, *not* the prescriptivist middle-school sense) of using the word “fun” as an adjective, in addition to its standard use as a noun. Anyone is perfectly likely to use *fun* as a noun (“I had a lot of fun at the party”); in addition, many of those people, though not all, will also be comfortable saying “It was a really fun party,” with an adverb modifying *fun*. However, the older generation of speakers would not say, “It was the funnest party I’ve been to this year,” while younger folks will find this adjectival use (note the –est adjectival suffix) perfectly acceptable. Another youthful example of a standard noun being called into use as an adjective as language changes is “dope,” meaning (we gather, since we are too old to find ourselves using this) “cool” or “hip.” A case of an adjective being used as a noun is the currently popular phrase “my bad.” As you can see, language is constantly in flux across groups of people and even to some degree in individuals over their lifetime. As psycholinguists we examine the competence of any particular individual speaker, at some point in time. We investigate the knowledge that that speaker has that allows him or her to speak and understand.

So What is Language Anyway?

We have stated that all and only humans use language. Animals do use communication systems, some rather sophisticated. These systems, however, do not exhibit all the hallmark properties shown by language. We will want to consider what are the characteristics common to all communication systems and what characteristics might be exhibited by language alone. These characteristics were generally enumerated by a linguist named Charles Hockett in the 1960s and have been presented over the past half century in various versions. Here are the properties we think are important for you to know.

All communication systems, by definition, must have a means of transmitting a message. Humans generally use a vocal-auditory **mode** for their language, but a manual-visual mode is also possible, as can be found in signed languages. No human language, however, incorporates whistles, foot stamps, or claps (though other nonlanguage types of human communication may use these). Some animals use a hormonal-olfactory chemical mode of transmission to communicate group recognition, alarm, sex, territory, or aggression information. For example, moths convey sexual information via pheromones, while cockroaches convey aggression, and hyenas convey territory marking. Also, the signals of all communication

systems, again by definition, must be **meaningful**; they are not random or non-sensical. The signaling in a communication system also serves a useful function for the animal in its environment, that is, it is **ecological**. It may aid in finding food or finding a mate or protecting offspring, for example.

Two other properties of communication systems have to do specifically with the relation between individual communicators. Some animal communication systems, including human language, must be learned through **interaction** with other individuals sharing that communication system. All babies can and will learn whatever language (or languages) they are sufficiently exposed to in childhood through interaction, regardless of the language spoken by the biological parents who contributed their DNA. In humans all aspects of the language system require exposure to other individuals to develop successfully, even though the ability to learn language is an innate genetic endowment. In some other species, such as in some birds, certain aspects of the communication system are learned while other parts are genetically coded. In yet other species, such as some insects, all communication is encoded directly in the genetics, so deviations, modifications, or innovations to the system by the creature simply are not possible. Humans, and a number of sophisticated animal communicators, also exhibit **reciprocity** in their communications, meaning that any particular individual can both create a communicative signal and understand such a signal. I can speak an utterance to communicate a message and also I can understand such an utterance if someone else speaks it to me. Some animal communicators may only send one type of signal, for example, indicating their sex, and another individual may only be able to receive that signal. A female organism in such a system can perceive the "I am male" signal but can't send it, and vice versa. Such a communication system does not exhibit the feature of reciprocity.

Some communication systems, including human language, have the property of **arbitrariness**. This means that the form of the signal (e.g., a word form) is not required to be related to (e.g., sound like) the thing it represents. In fact, it rarely does (which is to say that onomatopoeia is the

The mosquito species that carries yellow fever communicates mating availability by an acoustic matching of the sound created by its wingbeats. While the male wingbeat sound is normally higher in pitch than the females, when near one another these mosquitoes signal mating availability by converging to an even higher tone that is an overtone or harmonic of these two frequencies. Thus the **mode** could be said to be wingbeat-auditory (though their auditory system is rather different than a mammal's); the signal is **meaningful** and **ecological** in that it conveys information about an important activity for the animal in its environment – the opportunity for reproduction.

exception, not the rule, and even onomatopoetic words vary arbitrarily from language to language). Many animal communication systems use iconic signals and therefore do not exhibit the property of arbitrariness. Honey bees use movements that indicate the direction to a food source by effectively “pointing” to it; distance to the food is also related directly to the speed of the bee’s dance; and this appears to be an innate, unlearned, behavior of honey bees. We humans, however, must learn the words of our particular language; there is nothing that requires that some particular sequence of sounds have some particular meaning, as we saw with the *eagle/hedgehog* example earlier. Further, the allowed patterns found in a particular language for how words may be combined is also arbitrary and language-specific. There is nothing in the environment that requires that an adjective precede or follow the noun it modifies; different languages make an arbitrary choice.

Scientists studying language have arrived at further features that, linguists would argue, distinguish human language from all other forms of natural communication. The first unique feature of language is that messages are generated from recombining parts – this is the property of **compositionality**. The recombining parts that compose messages are called the language’s **discrete units** and crucially the patterns of recombination are meaningful; they encode meaning differences. Messages in human language are communicated by phrases composed of words. Compositionality also applies at the level of words. Words are composed of combinations of sounds or vocal tract actions that individually do not have any meaning of their own. There are a limited set of speech units in any particular language that can be recombined to form meaningful words. Messages communicated by nonhuman animals generally do not show compositionality. Nonhuman animals use signals that cannot be broken down into parts that the animal can recombine to make new or alternate messages. The signal communicating the message is, if modulated at all, modulated in a continuous way – say for example, in loudness or strength or rate. It is true that scientists have argued about the existence of compositionality in animal communication from time to time, but overall naturally occurring animal communication systems appear to lack compositionality.

Another particular property of human language is that it incorporates the possibility of communicating about objects, events, and emotion that are not in our immediate environment – a dream you had last year, a graduation ceremony you hope to participate in next year, or the surface of Neptune. Linguists have called this language property **displacement** because messages can concern items displaced in space and time. Messages are not necessarily driven by stimuli present in the animal’s – human’s – environs. Human language even allows us to talk about things that patently don’t exist, like a dessert you’ve never had but imagine tasting

or the wedding you called off. It is reasonable to attribute this ability to the complexity of human cognition, but even acknowledging that, human language still provides the means for communicating about the past, the future, and counterfactual situations.

Human language also has the potential to express an infinity of messages – it is an **open-ended** system. There are infinitely many sentences possible in a language, and new words can always be made up and added to a language when there is a use for them; indeed, this happens all the time. This may seem an obvious property of human language, but it has an important implication that you may not have thought of. The fact that all human languages are **open-ended** means that humans cannot learn their language by memorizing a set of possible messages. Other animals appear to learn (or innately be genetically provided with) a fixed set of messages that they use to communicate. Unlike these other animals, you will always be able to say something new that you have never said before and even that you have never heard said before. Try it for fun right now: think up a sentence that you believe you have never, ever said and never, ever heard before. It's not even a very hard challenge for us humans. For other animals, such a communication is not possible.

What makes possible this productive capacity of human language? Linguists sometimes use the term **generative** to describe language's combined properties of **compositionality** and **open-endedness**. Language is **generative** because it is based on a systematic relation of meaning and sound created by body actions. Meaning is of course internal; for meaning to become accessible between a speaker and listener – between a language producer and perceiver – it must be transmitted in the environment. By this, we simply mean that humans do not have ESP! Remember the speech chain we looked at above? This transmission is done by combining units of production (which we will discuss at length in later chapters) that can be recovered and decoded by the listener.

Linguists are interested in understanding the cognitive units used as the building blocks of human language and the system of relations among units of different sorts. An important system in language is its **syntax**. Syntax is the system of how words may be arranged in an utterance to convey meaningful relations among them. Consider, for example, the sentences:

Toby gave the book to Dani.
Dani gave the book to Toby.

The syntax of English determines that in the first sentence it was Toby who did the giving and that what he gave was a book and the person to whom he gave it was Dani. The second sentence specifies a different set of relations between Dani and Toby. Syntax refers to the structuring of units of meaning (for now, words) in sequence via structural (hierarchical)

Table 1.1 Properties of Communication Systems

Meaningfulness	All communication systems	↓
Ecological validity		
Reciprocity	Some communication systems	
Learned through interaction		↓
Arbitrariness		
Compositionality	Viewed as properties particular to human language	↓
Displacement		
Open-endedness		
Duality of patterning		↓

organization or relations between words. (We will discuss the notion of hierarchical organization in later chapters.)

Words are meaningful units and are part of an expandable set. But words also are composed of cognitive units. Linguists call the system that governs the organization of units composing words **phonology**. Phonology structures a relatively small set of units that is *not* expandable. In English, for example, the sounds composing “top” can be rearranged to form “opt” or “pot” (though not “pto” or “otp”), but a speaker of English could not wake up one morning and decide that a brand new sound, let’s write it <!>, could be used to form words in English. In any particular language, words, new or old, must draw from a stable, small set of nonmeaningful units called **phonological units**. *So in human language the meaningful messages (both sentences and words) are infinite in variety by virtue of the fact that words are produced from a system of combining a finite set of meaningless units.* Linguists, since Hockett in the 1960s, have described this hallmark property of language as **duality of patterning**.

What Scientists of Language – Including Linguists – Don’t Do (at Least for a Living)

Scientists who study language are not interested in prescriptive grammar, that is, rules that some authority decrees ought to be followed in speaking and/or writing a language – things like “don’t end a sentence with a preposition” and “don’t split an infinitive” and “don’t say *ain’t*.” (Notice how often the folks who apparently hold these positions of authority come up with rules starting with “don’t.”) Many of these rules in fact come from historical idiosyncrasies of a language. Prescriptivist views of language are often motivated sociologically, as nonstandard dialects of a language are often held in poor regard. Because of this, using these dialects can