University of California

Los Angeles

Making Words

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Philosophy

by

Ashley L Feinsinger

2016

Abstract of the Dissertation

Making Words

by

Ashley L Feinsinger

Doctor of Philosophy in Philosophy

University of California, Los Angeles, 2016

Professor Samuel John Cumming, Chair

There is a view implicit in of much of the recent work in philosophy of language according to which public words provide a reliable mechanism for communication. It is often assumed that two linguistic agents can come to understand one another to the extent that (i) they both use the same public words with the same meanings, and (ii) they can recognize which of those words are being used. But despite this central role in communication, serious investigations into the nature of words are lacking. How should we understand these objects and how are they related to their utterances? In virtue of what are two utterances utterances of the same word?

Providing answers to such questions is complicated by the fact that there are conflicting demands on accounts of public words as they function in communication. While traditional theories of communication require that agents use the same words and can recognize which words are being used, it is a common experience that the same word can be pronounced in a variety of ways. It is not obvious what kind of model can reconcile the existence of such variation with the fact that communication succeeds despite it. Furthermore, it's not obvious what a public word *is* if different agents may pronounce it differently. But without such an

account, many philosophical theories of linguistic communication are incomplete.

In *Making Words*, I propose a novel account of public words that is responsive both to their communicative role and the facts of pronunciation variation. I argue that the relation between a public word and its articulations is grounded in a particular kind of social relation that holds between the speaker and those spoken to. On the model offered here, using the same word requires not that agents make the same sounds, but rather that agents coordinate their internal linguistic representations. This coordination, I'll argue, is achieved by the existence of communicative policies that agents establish with one another in service to their joint goal of communication.

By thinking of public words as the products of these communicative policies, I argue that we can recover an account on which agents may use the same words, pronounce them differently, and track which words are being used in a communicative exchange.

This dissertation asks that we abandon the orthodox idea that public words are stagnant objects with fixed pronunciations in favor of a new idea: that public words are dynamic objects, made by the speakers and receivers who use them and related to their articulations via the social relations that hold in a linguistic community. Indeed, it argues that where there are no social networks of communicative policies, there can be no notion of a public word as a vehicle of communication.

The dissertation of Ashley L Feinsinger is approved.

Joshua Armstrong

Kent Johnson

David Kaplan

Samuel John Cumming, Committee Chair

University of California, Los Angeles

2016

For Daniel

Table of Contents

List of Figures

Acknowledgments

First and foremost, thanks to my advisor, Samuel Cumming, whose intellectual scrutiny and unwavering encouragement have undoubtedly made this work possible. From Sam I've learned many things, including how to be a philosopher, how to scrutinize an idea without losing sight of the larger dialectic, and the worth of both philosophical formalization and philosophical prose. Additionally, his own work has had a profound influence on this dissertation and on my intellectual maturity, and I owe most of what's good in this dissertation to him. But by far, the most valuable thing I've learned from Sam is the incomparable importance of being inspired. I never left one of our meetings without a renewed sense of excitement about my project and about philosophy in general. He always reminded me, by impeccable exemplification, of the beauty and power of philosophy. That is no small thing, and I've never taken it for granted. It has made an indelible difference in my life.

My interactions with David Kaplan began the first year of my undergraduate education. I credit David entirely with sparking my interests in words, philosophy of language, and philosophy more broadly. I've learned vastly from his overwhelming intellect and stamina, as well as his sometimes mystical insight into the true nature of language. David's own work on words showed me that this was a rich and important area of study, and this dissertation benefited greatly from his relentless but encouraging questions, criticisms, and counter examples. I have personally benefited from his friendship and guidance, and I often reflect on how lucky I've been to cross his path.

I began taking the ideas in this dissertation seriously after one of Josh Armstrong's talks at UCLA, and much of his own insights have made their way into

my own work. His vast knowledge of the literature and his deep understanding of the complicated ways in which philosophy of language, linguistics, and cognitive science work together has undoubtedly made this work better. It has also changed how I do philosophy. He was always willing and able to talk about anything at anytime, and his unparalleled enthusiasm and positivity have kept philosophy fun at times when it would have otherwise been, well, not fun.

To Mark Greenberg and Kent Johnson, thank you for providing your expertise and for reading drafts even in the eleventh hour. And thanks to Jessica Rett for being my window into theoretical linguistics, for educating a naïve philosopher in formal semantics, and for providing a linguist's perspective on the issues in this dissertation. I'm grateful for the patience and kindness you've shown me in our meetings, and for the unofficial mentor-ship I conned you into.

I've also been fortunate to have Tyler Burge as a mentor and friend during my time at UCLA. We began our conversations in an independent study my junior year as an undergraduate, and I've learned quite a lot about philosophy, teaching, and life from our ongoing friendship. I've always felt I've had someone in my corner.

This work has grown immensely as the result of the rich intellectual community at UCLA, including those in the Philosophy of Language Workshop, the Philosophy of Language Seminar, and the Syntax-Semantics Seminar. Special thanks goes to Greg Antill, Denis Buehler, Lee-Ann Chae, Lindsey Chambers, Gabriel Greenberg, Pamela Hieronymi, Andrew Hsu, Brian Hutler, Andrew Jewell, Kevin Lande, Elliot Michaelson, Paul Nichols, Terry Parsons, and Alex Radulescu. I owe a particular debt to Tamar Weber, who on many occasions and without hesitation, has offered her company, apartment, and conversation. I've benefited from our many work dates and conversations about concepts, language acquisition, and

(to her dismay) Paderewski. Our intellectual and personal relationship has been invaluable.

Thank you to my parents, Suzanne and Paul Feinsinger, for teaching me the value of both a good argument and a good work of art, for unequivocally supporting me financially, intellectually, and emotionally, and for never asking me what I was going to do with a philosophy major. To my sister, thank you for always making me laugh. Our house is always filled with music, laughter, and deep discussion—I wouldn't have gotten anywhere without those things.

Lastly, my greatest debts are to Daniel and Liv Aharoni, the tallest and smallest of word-makers. To Daniel, thank you for seeing in me a version of myself that I would be lucky to embody and for your unsurpassed perseverance and kindness. To Liv, thank you for being a constant source of light in a world of dim.

Vita

2006	B.A. (Philosophy with a minor in Linguistics), University of California, Los Angeles
2008	M.A. (Philosophy), University of California, Los Angeles
2011	C.Phil (Philosophy), University of California, Los Angeles

Publications and Presentations

"The Metaphysical and Social Foundations of Public Words", UCLA Dissertation Workshop, 2015

"Existing, Possible, and Innovative Words", California Universities Semantics and Pragmatics Conference at University of California, San Diego, 2012

"Existing, Possible, and Innovative Words", UCLA Syntax and Semantics Seminar, 2012

"The Semantic Lexicon and Analyticity", UCLA Philosophy of Language Workshop, 2011

"On Words", UCLA Philosophy of Language Workshop, 2010

"The Semantic Lexicon and Analyticity", *UCLA Working Papers in Linguistics, Papers in Semantics*, 2012, Volume 16, pp. 51-68.

CHAPTER 1

Introduction

> "God, having designed man to be a sociable creature, not only made him with an inclination and a need to have fellowship with other men, but also equipped him with language, which was to be the great instrument and common tie of society. So nature shaped man's organs so that he could make articulate sounds, which we call 'words'."
>
> ---
>
> John Locke, *An Essay Concerning Human Understanding*, 1689.

Consider the following scenario. You and I have been set up on a blind date, and we need to decide where to go and when. I call you on the phone, suggest we see a movie at a local theater, and tell you the various options and show times. You respond with your choice, and we both show up at the same time and place and see the movie.

This scenario is fairly mundane. Indeed, it and exchanges like it happen all the time. But how is it, really, that after I made some noises by pushing air through my vocal tract, and you made some more noises by pushing air through your vocal tract, we successfully met each other at the movie? How do we, by making certain sounds, come to understand one another? In the scenario above, you and I hadn't previously spoken with one another, we hadn't previously agreed on what our noises were to mean or refer to, and we might have even made different noises to mean or refer to the same things. (Suppose I said, "Shall we meet at *ee-leven*

a.m.?" and you replied with, "Sure, see you at *uh-leven* a.m.")

Often, philosophical answers to these questions involve the notion of a *word*, and appeal to the idea that some of these noises count as *utterances* of words. On these views, you can make yourself understood by uttering some of these words, and I can recover the intended meaning of your utterance by first recognizing the words that you uttered. Notice that on this story, you and I can use the *same* words. Words are assumed to be public sharable entities.

But despite this central role in communication, philosophers of language have been surprisingly quiet on the nature of words and their metaphysical properties. Perhaps this is because some have felt that words and expressions are ontologically safe, or at least more ontologically safe than the meanings they express.[1] Indeed, while meanings have been called "creatures of darkness" [37], words have earned no such epithet. And, perhaps it is for this reason that much of the philosophical literature on words concerns the notion of meaning, while little attention is given to words as signs themselves—as things spoken and uttered. But as this dissertation will suggest, words are mysterious objects, and we have reason to be wary of both their apparent metaphysical security and the lack of scrutiny they've received.[2]

What are words? How are they related to their utterances? What does it take for two agents to use the same word, or for two noises to count as utterances of the same word?

This dissertation fills this void in philosophical discussions by exploring the nature of linguistic expressions—things like words and sentences—as objects that

[1]For example, we find such sentiments in Kaplan [26] when he says, "The ontological status of meanings or senses is less well settled than that of expressions", and, "Now what shall the sentences denote? For my present purposes it will suffice to take the ontologically secure position and let them denote expressions...". Also, see Quine [37].

[2]This concern is echoed in the more recent work of Kaplan [27] and [28], who, along with Hawthorne and Lepore [23] and Cappelen [6], seriously take up metaphysical questions about words.

are *uttered* or *articulated*. More specifically, it aims to fill this philosophical void in a way that is informed by work in empirical linguistics (particularly contemporary phonology, as it has been developed under Noam Chomsky's influence), and work in sociolinguistics on the social factors affecting patterns of word articulation.

I begin by laying out what public words must be like in order to play the role philosophers want them to play in enabling linguistic communication. I'll then present a challenge: that any account of words in their communicative role must accommodate the existence of pronunciation variation. People speak in different ways, and it's not obvious how agents communicate in spite of this. Furthermore, it's not obvious what the public word is if speakers can pronounce it differently. The existence of such variation has been taken to threaten both the philosopher's theory of communication as well as the philosopher's notion of a word.

I'll offer a new view of public words, on which the relation between a word and its utterances is grounded in a kind of social relation that holds between the speaker and those spoken to. I'll see the notion of a public word and the same-word relation as grounded in the fact that agents may coordinate their internal linguistic representations in the service to their joint goal of communication.

The seeds of this view are contained in the Lewisian notions of convention *coordination*, according to which agents align their actions or plans of action with one another in an effort to satisfy some mutual goal. Agents who seek to communicate via the use of a linguistic signal must coordinate their strategies for speaking and understanding, and I'll argue that we should understand public words by first thinking about this joint activity.

On the view developed here, what I will call *The Communicative Policy Model*, using the same word does not require agents to make the same sounds. Instead, it requires that those agents interpret each other's utterances in ways that pair up

their internal linguistic representations. I'll suggest that we may build the notion of a public word out of sets of these coordinated representations, recovering a notion of public word that not only fits philosophical models of communication, but also accommodates pronunciation variation. On such a picture, pronunciation variation poses no special challenge in accounting for public words and communication, since agents may coordinate their representations despite pronouncing them differently.

A public word will turn out to be more like a social contract or convention than a fixed sequence of sounds. This is because, as I will argue, the conditions under which we have a public word (or under which two agents are using the same word) depend on the strategies or communicative policies that agents establish with one another. Just as we might say there is a social convention among a group of agents when there exists jointly accepted principles of action as to how one is to behave in certain circumstances, we may say there is a public word in use among a group of agents when there exist jointly accepted principles of action as to how one is to behave linguistically in a certain circumstance.

Furthermore, just as a social convention might involve a complicated course of action in which agents will do different things (e.g., I will cook the food and you will wash the dishes), using a public word might also turn out to involve a complicated course of action in which agents will do different things (e.g., I say "ee-leven" and you say "uh-leven").

The account I offer is importantly different from previous models in the following ways. First, it is informed by the empirical claims and predictions of phonological theory (as developed in the Chomskyan tradition) and by the sociolinguistic data concerning patterns of pronunciation variation. Broadly speaking, current inquiries into natural language are in one of two traditions: the Chom-

skyan generative linguistic tradition, which seeks to understand the psychological mechanisms devoted to forming linguistic structural descriptions, or the philosophical tradition, which seeks to understand linguistic communication. While the aims of these traditions are often orthogonal (indeed, Chomsky denies that the aim of linguistic theory is to explain communication and rejects the very notion of a public language used for this purpose), my account integrates both traditions. It brings together the mechanisms posited by theoretical linguistics with the manifest communicative function of language and advances the hypothesis that a complete understanding of natural language cannot be achieved otherwise.

Phonologists have a lot to tell us about how agents represent and systematically manipulate linguistic sounds, and philosophers have a lot to tell us about how groups of agents use words to understand one another. But neither has had much to say about how the sounds an individual makes are related to the public words she uses, and it's my contention that neither field can give an adequate answer on its own.

Second, the view advanced here makes a fundamental appeal to the audience in coming to understand the relation between a word and its utterances. I'll argue that previous models of words have failed to account for the pronunciation variation facts because they appeal solely to facts about the speaker in coming to understand which word an utterance is of. But on my view, the audience plays an important role in two ways. First, a speaker's pronunciations can vary depending on features of her audience, and second, which word a speaker is uttering importantly depends in part on how her audience plans to interpret her utterance.

I'll argue that facts about which words a speaker has uttered and whether two articulations belong to the same word depend just as much on how the audience plans to interpret those utterances as they do on how the speakers plans to produce

them. This dissertation advocates for the position that we cannot fully understand the same-word relation and the notion of a public word by thinking solely about the speaker and the sounds she produces—we must also look to who is being spoken to.

I won't be offering a theory of public words, in the sense that I won't be offering strict necessary and sufficient conditions for individuating public words or for saying when an utterance is an utterance of a particular word. Instead, I'll be advocating for a kind of theory, on which the notion of a public word is built out of the communicative policies that hold between agents in a linguistic community. However, in advocating for this kind of theory, I will offer a candidate necessary condition on two utterances being utterances of the same word and on an utterance being a utterance of a particular public word.

This dissertation asks that we abandon the orthodox idea that public words are stagnant objects with fixed pronunciations, living in Plato's heaven. It offers instead a new idea: that public words are dynamic objects, made by the speakers and receivers who use them and related to their utterances via the social relations that hold in a linguistic community. It argues that where there are no social networks of communicative policies, there can be no notion of a public word as a vehicle of communication.

The dissertation will be organized as follows. Chapter 2 lays out the desiderata on a theory of public words as they function in communication and suggests that satisfying them won't be easy. In particular, it focuses on the tension between the existence of pronunciation variation and the following two things: (i) systematic, reliable, linguistic communication and (ii) the notion of a public shareable word. While Davidson has taken the existence of linguistic variation to tell against traditional theories of communication, and Chomsky has used variation to argue

against a substantive notion of a public word, I offer a different hypothesis: that linguistic variation suggests we need to *change* our notion of a public word. We can maintain that public words play a robust role in communication, that pronunciation variation is pervasive, and that agents can reliably communicate in the face of linguistic variation, by thinking of words in a different way.

Chapter 3 presents previous models of words and argues that they fail to meet our desiderata. It surveys *Replication Models*, on which the same-word relation is determined by a match in phonological shape, Kaplanian *Intentional Models*, on which the same-word relation is determined by a speaker's intentions, and Chomskyan *Individual-Lexical Models*, on which the same-word relation is determined by a speaker's internal phonological grammar.

I'll argue that *Replication Models* fail to account for pronunciation variation, *Intentional Models* fail to adequately constrain pronunciation variation, and *Internal-Lexical Models* fail to account for social pronunciation variations and lack a notion of a public sharable word. I'll highlight the fact that these views struggle for the same reason: they ignore the audience and her interpretive strategies in coming to understand the metaphysical facts about words.

Chapter 4 introduces *The Communicative Policy Model*, on which the same-word relation (and subsequently the notion of a public sharable word) is based in sets of overlapping communicative policies adopted by both speakers and hearers in service to their joint goal of communication. On this view, using the same word requires agents to coordinate a particular pair of internal linguistic representations. These policies, as I'll conceive of them, are akin to Lewis' signaling strategies, and they constitute solutions to a coordination problem posed by linguistic communication.

I'll argue that in order for utterances made by distinct agents to be utterances

of the same word (i) those agents must have communicative policies that map those lexical items onto one another, and (ii) in making their utterances, those agents must be acting in accordance with those communicative policies. On this model, agents can pronounce the same word differently (as predicted by generative phonologists and sociolinguists), agents can track which pronunciations belong to which words (as communication requires), and agents can cope with pronunciation variations and words they have not previously encountered.

Chapter 5 asks and answers various questions about the view offered in Chapter 4, including how *The Communicative Policy Model* might individuate public words, what it says about the possibility of pronunciation errors, how agents might be coordinated with other agents indirectly by appealing to the policies at work in the wider linguistic community, and how to deal with Kripke's famous Peter and his confusions about Paderewski.

CHAPTER 2

Words and Communication

2.1 Introduction

There is a view implicit in much of the recent work in philosophy of language according to which public words play a central role in enabling linguistic communication. It is often assumed that linguistic agents can come to understand one another to the extent that (i) they use the same words and associate the same meanings with those words, and (ii) they can track which of these words is being used. In fact, it is sometimes argued that the only way to explain the ease and reliability of linguistic communication is to posit that agents use the same public words.

This chapter is about two things. First, it is about philosophical theories of linguistic communication and what words must be like in order to enable such communication. Second, it is about a challenge brought against these theories of communication as well as the notion of a word they assume.

I'll begin by outlining how, on many traditional views of linguistic communication, words provide a reliable mechanism for the transfer of information. They are able to fill this role in part because they are public, shareable objects.

Next, I'll outline the well-known and well-studied fact that agents speak in different ways—that pronunciation variation is ubiquitous. And I'll claim that any account of words in their communicative role must accommodate the existence of

pronunciation variation.

But the existence of linguistic variation has been taken by some to challenge both the philosophical picture of linguistic communication and the philosophical notion of a public word. Some philosophers, such as Donald Davidson, have taken the existence of linguistic variation to threaten the utility of public words in explaining communication, and some linguists, such as Noam Chomsky, have taken linguistic variation to threaten the possibility of public words altogether. I'll rehearse these objections to the standard picture of communication and public words, concluding that they rightly show that something about these pictures is in trouble.

But I'll reject the claim that in the face of variation, we must give up traditional theories of linguistic communication and the notion of a public word altogether. Instead, I'll suggest that we need a new notion of a public word.

2.2 The Communicative Function of Words

Whenever agents aim to communicate, they face a problem. They have to find a means of outwardly expressing their inner thoughts, desires, goals, etc. In the case of linguistic communication, agents have to decide which sounds to make in order to successfully communicate those thoughts. And, each agent's choice of sounds depends on her expectations about other agents: agents want to make sounds that their audiences will understand.

Agents solve this problem all the time. Furthermore, they solve it in reliable ways. They solve it in ways that will continue to be successful in the future, and they solve it in ways such that the speaker can *expect* to be understood in making her utterance and the interpreter can *expect* to correctly understand in interpreting those utterances. How do agents do this?

Many standard philosophical theories of linguistic communication are versions of what I will call *information-transfer* theories. On such theories, a linguistic communicative exchange is successful when the interpreter entertains the content the speaker intends her to entertain on the basis of receiving the utterance and understanding what it was intended to express. Such theories require a reliable mechanism for matching content, and many of them assign a central role to public words in providing such a mechanism.[1]

For example, suppose you want me to entertain the proposition that Hesperus is beautiful, and so you utter something that sounds very much like, "Hesperus is beautiful". I can reliably uncover the intended meaning of your utterance, the story goes, if (i) we both use the same language, which in part consists of using the same public words with the same meanings, and (ii) I can recognize which words you've uttered. Indeed, we would not expect our exchange to succeed were I to assign the meaning *Saturn* to the word 'Hesperus' or think that you articulated the word 'Jupiter'.

Notice that not all methods by which I can come to entertain the intended content are equally good, as the following three cases suggest.

(A) Suppose that you intend for me to entertain the content that Hesperus is beautiful, and to this end you utter, "Hesperus is beautiful". But someone covers my ears (or knocks you out before you are able to utter those sounds), and then hits me over the head. As a result of the blow, I spontaneously come to entertain the content that Hesperus is beautiful.

(B) Suppose that you intend for me to entertain the content that Hesperus is

[1] I'll talk primarily about communication in terms of matching content, but even views that don't conceive of information transfer as content matching still require a reliable mechanism for mapping sender states onto the appropriate receiver states. See Buchanan [4] and [5] for a nice discussion of what he calls "the standard view" and arguments against it.

beautiful, and to this end you utter, "Hesperus is beautiful". I hear you, and I correctly identify what words youve uttered. But, I assign an interpretation to your utterance by rolling a die. As luck would have it, I come to entertain the content that Hesperus is beautiful as the result of the die roll.

(C) Suppose that you intend for me to entertain the content that Hesperus is beautiful, and to this end you utter, "Hesperus is beautiful". But I wasn't paying very much attention to your utterance, and as a result, I am unable to identify the first word you've said. I guess correctly that it was the word 'Hesperus' by picking a name out of a hat and come to entertain the content that Hesperus is beautiful.

The cases in (A), (B), and (C) seem to be defective linguistic communicative exchanges, and they highlight that the content matching process matters. The case in (A) is defective despite the fact that content was matched, and the cases in (B) and (C) are defective despite the fact that the words were correctly identified and content was matched.

The problem in all of these cases seems to be that the content matching process, although successful, is *unreliable* or *lucky,* in way that content matching in successful linguistic communication is not.

In contrast, if I proceed via my knowledge of semantic conventions (which associate public words with their meanings) in conjunction with my ability to recognize which words have been uttered, this method can be reliable in ways that other methods are not. *Reliable* communication can be achieved if agents use the same words with the same meanings and those agents recognize which words are used. Any adequate theory of public words that aims to be responsive to philosophical accounts of linguistic communication should thus meet these desiderata.

2.3 Pronunciation Variation

Any adequate model of public words in the communicative context must account for the particular ways in which a word's utterances can vary in pronunciation. That is, it must account for what I will call *Tolerance* and *Constraint*:

Tolerance: the utterances of a word can vary in their phonological features

Constraint: the variation described by *Tolerance* is constrained

For example, the word 'water'[2] is pronounced by some American English speakers with an intervocalic flap (e.g., [wɔɾəˬ]),[3] by some British English speakers with the voiceless stop 't' (e.g., [wɔtəˬ]), and by other British English speakers with a glottal stop (e.g., [wɔʔəˬ]).[4] Some English speakers pronounce the word 'pen' with a higher vowel (e.g., [pɪn]) than other English speakers (e.g., [pɛn]),[5] and some English speakers articulate the word 'cat' with a raised diphthong (e.g., [kɪət]), while others do not (e.g., [kæt]).[6]

Some New Yorkers fully pronounce consonantal 'r' in post-vocalic position (e.g., articulating the word 'floor' as [flɔr]) while others do not (e.g., articulating the word 'floor' as [flɔ]), and some residents of Martha's Vineyard centralize their articulations of the vowels [aʊ] and [aɪ] (e.g., articulating them as [əʊ] and [əɪ]) while others do not [31]. Some speakers omit the word initial 'h' in articulations

[2]I'll refer to public words by using their standard English spellings enclosed in single quotes, and I'll refer to the phonetic form or pronunciations of utterances by using IPA transcriptions enclosed in square brackets.

[3]This alternation is often called "tapping" or "flapping". See Hayes [24, p.32].

[4]This alternation is often called "t-glottalization". See Bergmann et al. [2, pp.102-103]. See chapter 3.2 for a list of the kinds of phonological alternations in English.

[5]This variation is known as the "pen-pin merger"[2, p.423].

[6]This variation is knows as the "Northern Cities Vowel Shift" [2, p.422].

of 'here' (e.g., [ɪr]) in some contexts but not others (e.g., [hɪr]).[7]

Importantly, while the phonological features of a word's utterances can vary, it is crucial that not anything goes. While some variations are tolerated and support communication, others are not and do not. Just as a performance that sounded exactly like a performance of the song 'Hey Jude' does not count as a performance of the song 'Let it Be', an utterance with the pronunciation [kæt] does not count as an utterance of the word 'dog' [23]. Mere grunts don't count as utterances of any word. This is to say that there are constraints with respect to how a particular word can be articulated. Moreover, these constraints shift: [pɪn] might be an utterance of the word 'pen' in some contexts but not others (depending on who is speaking). Such constraints seem to serve the communicative function of putting pressure towards a one to one mapping between meanings and forms.

2.4 The Problem of Variation

The existence of linguistic variation is sometimes taken to threaten, to various degrees, aspects of traditional theories of linguistic communication and the notion of a public word they assume.

Traditional theories of linguistic communication claim that communicative success is due to the fact that agents have prior shared knowledge of a language. Perhaps the most prominent theory of a public language is that of David Lewis, on which a language is a set of sentence-meaning pairs used as matter of convention by a population. On such a picture, communication is said to be reliable among a group of agents to the extent that those agents share knowledge of these conventions [33].

[7] This alternation is described as 'h-deletion' and can apply in rapid speech to unstressed syllables (e.g., "Can you give her his hat" becomes: [kæn ju gɪv ər ɪz hæt].) See Bergmann et al. [2, p.114].

In the same spirit, we might conceive of a public word as a pronunciation-meaning pair used by a group of agents as a matter of convention.[8] Or, on a slightly different formulation, a public word is an expression used by a group of agents with a conventionally established pronunciation and a conventionally established meaning. On such a picture, we might say that an utterance is an utterance of a particular word only when it manifests that expression's conventionally established pronunciation.

A model of words based in the Lewisian notion of convention might account for some of our desiderata rather straightforwardly: (i) words would be public and usable by more than a single agent, and (ii) communication via the use of words would be reliable (and agents would be able to track which words are being uttered) to the extent that they have knowledge of and act in accordance with the conventions which determine a word's meaning and its pronunciations.

Unfortunately, this view of communication and the words that enable it is complicated by the existence of linguistic variation. In particular, these notions are complicated the fact that there exists pronunciation variation (of the kinds discussed in the previous section) among groups of agents said to be sharing and using the same public words.[9]

Some have taken the existence of linguistic variation (i) to tell against the very utility of conventions (and thus the role of these Lewisian public words) in enabling communication and (ii) to undermine the possibility of public words altogether. I'll discuss each of these cases in turn, first addressing Davidsonian objections to

[8] Armstrong adopts a similar notion of an expression in a public language [1].

[9] We might try modifying the Lewisian view on which a public word is conventionally related to a *single* pronunciation, allowing that a public word is conventionally related to multiple pronunciations. By requiring that agents have knowledge of multiple pronunciation conventions, we might preserve the traditional explanation of communication while admitting pronunciation variation. This is, in spirit, the kind of view I will advocate for. However, it will require changing our understanding of the pronunciation conventions that underwrite linguistic communication.

traditional views of communication in the face of linguistic variation, and then turning to Chomskyan objections to the notion of a public word in the face of a heterogeneous speech community.

2.4.1 Threats to Communication

As Davidson points out, sometimes communication succeeds despite the fact that conventional interpretation of an utterance doesn't yield the intended meaning. Furthermore the agents seem to have no prior shared knowledge of a convention that would assign the utterance its intended meaning [17]. Some examples involve what Davidson calls "malapropisms". A few examples are given below:

(i) Yogi Berra: "Texas has a lot of electrical votes."

(ii) Reverend Spooner: "The lord is a shoving leopard."

(iii) Archi Bunker: "We need a few laughs to break up the monogamy."

(iv) My college roommate: "Dont worry, I can interpretate that for you!"

As Davidson argues, if communication succeeds in these cases, it can't have been via the agents' prior shared knowledge of linguistic conventions. Why not?

Consider (i), for example, and that the audience is able to move from the pronunciation [ʌlɛktrɪkəl] to the intended meaning *electoral*. How does the audience do this? One explanation is that audience succeeds by first recognizing that the pronunciation [ʌlɛktrɪkəl] in this context is a pronunciation of the word 'electoral'. Another possible explanation, which is the one Davidson opts for, is that the audience succeeds by recognizing that the speaker uttered the word 'electrical' and that this word (in this context) means *electoral*. But it is not likely that both the speaker and the audience have prior shared knowledge of either a pronunciation

convention specifying that [ʌlɛktrɪkəl] is a pronunciation of the word 'electoral' or a semantic convention specifying that the word 'electrical' means *electoral*.

What *is* likely is that the audience has knowledge of the convention that the pronunciation [ʌlɛktrɪkəl] is a pronunciation of the word 'electrical' and knowledge of the convention that the word 'electrical' means *electrical*. But were the audience to use these conventions in interpreting Yogi Berra, she would arrive at the content that Texas has a lot of electrical votes, which is not the content Yogi intended her to entertain. And so, we might conclude that the audience succeeds not via shared knowledge of linguistic conventions and not via the use of public words, but by some other means.

The challenge seems to generalize to any case in which communication succeeds despite the fact that agents don't share prior knowledge of linguistic conventions by which success would be possible. Notice that these cases need not involve "mistakes". The same problem arises each time an agent encounters a novel pronunciation. For example, the first time an agent encounters a Cockney British English speaker's pronunciation of the word 'water' (e.g., [wɔʔəˇ]), she doesn't yet have knowledge of a pronunciation convention associating that word with that pronunciation. But still, communication proceeds reliably and effectively. Such success, it is argued, can't have been via the prior shared knowledge of a pronunciation convention.

A similar worry arises whenever an agent encounters a novel word altogether. Agents coin new words all the time, weaving them effortlessly into conversation. You might ask me at breakfast if "the paper boy *porched* the newspaper yet" or at our annual holiday party, if I would *jellybean* the gingerbread house roof.[10] If communication succeeds in such cases, as it often does, it can't be via the use

[10]See Clark and Clark [12] and Clark [13] for a nice and thorough discussion of innovative words.

of our prior shared knowledge of linguistic conventions. I don't have knowledge of a convention associating the word 'jellybean' with the meaning *to cover with jellybeans* since arguably, before your utterance, there is no such convention in existence.

2.4.2 Threats to Public Words

> "[Public languages are] presupposed by virtually all work in philosophy of language and philosophical semantics ... This idea is completely foreign to the empirical study of language. Nor has anyone indicated what sense it might have as to how do we decide, for example, whether the word "disinterested" in the language I partially know is pronounced as in Boston or in Oxford, or whether it means uninterested, as almost all speakers believe (ignorantly, we are told), or unbiased, as certain authority figures insist? For the empirical study of language, the questions are meaningless. What are called languages or dialects in ordinary usage are complex amalgams determined by colors on maps, oceans, political institutions, and so on, with obscure normative-teleological aspects.In ordinary human life, we find all sorts of shifting communities and expectations, varying widely with individuals and groups, and no right answer as to how they should be selected." Chomsky [10, pp.18-19]

I take Chomsky to be making two claims against public words in the above passage. The first is that in the face of linguistic variation, there is no principled—or no principally linguistic—way to individuate public words and the communities that are said to use them. According to Chomsky, absent socio- and geo-political considerations, there is no non-arbitrary way to to decide who is speaking the

same language or using the same words. If Chomsky is right, and the notion of a publicly shared word is suspect, then so is the idea that agents succeed in linguistically communicating via the use of publicly shared words.

The second claim I take Chomsky to be making is this: the public objects assumed and invoked by philosophers are divorced from the subject matter of empirical linguistics, and therefore, of no theoretical interest. This is to say that even if we could find principled linguistic ways to individuate public words and their users, these words would play no role in explaining and understanding the kind of phenomena linguists are interested in explaining.[11]

We can see part of the motivation for these claims when we appreciate the goal of linguistic science. According to linguists working in the Chomksyan generative tradition, the goal of linguistic science is to uncover the psychological and biological mechanisms that underwrite an agent's idealized capacities for speaking and understanding, or her *linguistic competence* for constructing complex expression-meaning pairs from basic expression-meaning pairs. From this point of view, the differences or similarities between how two English speakers talk are no more interesting than the differences or similarities between how a Japanese speaker and a Spanish speaker talk. And, more importantly, nothing is gained in understanding the psychological life of a language user by seeing her as speaking the same language as another agent, or as speaking $English$ or $French$.

If Chomsky is correct, the very aims of this dissertation are problematic. I've set out to give a theory of words on which (i) they are public, (ii) they are used by groups of agents for the purposes of enabling linguistic communication, and (iii) they may enjoy the kind of pronunciation variation described by *Tolerance* and *Constraint*. But if Chomsky is right, the variation facts make it so that we won't

[11]The claims I take Chomsky to be making against public words are akin to the claims Armstrong [1] takes Chomsky to be making against public languages.

find the kind of words we are looking for (and that even if we did, such objects will be divorced from the subject matter of linguistic science and of no theoretical interest).

Even setting the Chomskyan motivations aside, we may still wonder, what *is* the public word if we all may pronounce it differently?

2.5 What to do?

Let's take stock. We looked at the role that public shareable words play on many theories of linguistic communication and I've claimed that any adequate notion will be one on which distinct agents can use the same words and pronounce them differently. We then rehearsed some objections to the possibility of fulfilling these desiderata. The first was that some pronunciation variation seemed to undermine the very picture of communication we are assuming, and the second was that linguistic variation in general undermines the notion of public shareable words (the very thing we are trying to understand).

Are we in search of a chimera? It is the aim of this dissertation to establish that we are not. I do not think that the ability of agents to cope with malapropisms, novel words, or novel pronunciation variations crowds out a role for public words in explaining linguistic communication, and I do not think that the existence of pronunciation variation threatens the possibility of public sharable words.

Instead, I take the Davidsonian and Chomskyan points to suggest that we need a new notion of a public word—one on which a word's pronunciations are not determined by similarity in phonological form, one on which agents need not have prior knowledge of the pronunciation conventions they seek to use in communication, and one on which the communities that are said to use such words are not carved up by merely socio- or geo-political considerations.

The rest of this dissertation is devoted to providing a way of understanding public words that meets all of our desiderata.

The first step towards coming to understand words in the communicative context is to change where we start our investigation. If we start from the idea that a word in an abstract object, which either is or has some fixed pronunciation, we face the challenge of how to specify the principles of variation. How far can two utterances be from one another in pronunciation and still count as utterances of the same word? Where is the boundary between a pronunciation of one word and a pronunciation of another?

As I'll argue in the next chapter, specifying the principles of variation is not straightforward. Which pronunciation variations a word enjoys is not simply matter of similarity in sound, or similarity in the kinds of features phonologists isolate. For example, the sounds [t], [k], and [ʔ] are all voiceless plosives (they are all produced without vibrating the vocal chords and they are all produced in by pushing air out of the oral cavity). The parts of the oral cavity at which [t] and [k] are made are closer together than the parts of the oral cavity at which [t] and [ʔ] are made. The [t] is articulated behind the front teeth, the [k] is articulated roughly in the back of the mouth, and the [ʔ] is articulated in the back of the throat. (As phonologists would say, the "places of articulation" for [t] and [k] are closer than the "places of articulation" of [t] and [ʔ].) But despite this fact, [wɔtə˞] and [wɔʔə˞] but not [wɔkə˞] are pronunciations of the word 'water'. These kinds of facts are hard to account for by taking a microscope to the sounds themselves.

Trying to draw boundaries between words and utterances on the basis of similarity in pronunciation might not be impossible, but I'll argue that it misses a more fundamental explanation about words and their utterances in communication. As I'll argue in the next chapters, we should start our investigation into

words by thinking first about what agents are doing when they are using them and what this activity is like. By understanding this activity, we can get a better understanding of how words may be pronounced in different ways, and how these pronunciations change and shift depending on various linguistic and social variables. When we zoom out and start looking at groups of agents in a linguistic community, we can recover a notion of a public word that will naturally accommodate the kinds of pronunciation variations described by generative phonology and sociolinguistics.

But before giving the details of this new account, I first turn to a detailed survey and criticism of previous models of words.

CHAPTER 3

Previous Models

3.1 Introduction

The goal of this chapter is to present and criticize three previous models of words and the relation between a word and its utterances. Although each of these theories is aimed at giving an account of slightly different, although related, phenomena, they each provide an answer to some of the questions surrounding the nature of words and the same-word relation.

The first of these theories are what I call *Replication Models*, which hold that a word is a kind of fixed abstract form, which all of its utterances instantiate. These theories primarily aim at telling us what a word is and then derive from this an account of when two utterances are utterances of the same word.

The second of these theories are what I call *Intentional Models*, on which a speaker's intentions determine which word an utterance is an utterance of. These theories say nothing about the nature of words themselves, but provide an account of the relation between an utterance and the word it is an utterance of. However, I will focus on a particular version of this model, what I will call *Historical-Intentional Models*, which do claim that words are a kind of naturalistic object, created by agents and individuated by their introduction events.

The third of these theories are what I call *Internal-Lexical Models*, on which a word is a kind of mental linguistic representation and on which a speaker's

phonological grammar determines which word an utterance is an utterance of. These views make two claims. The first is a reductionist claim that the only viable notion of a word is that of an internal linguistic representation in an individual's mental dictionary. The second is a claim about how those representations are articulated or expressed by an individual.

I'll argue that none of these views provide an adequate notion of a word and of the same-word relation in the communicative context. *Replication Models* fail to account for pronunciation variation and *Intentional Models* fail to account for the fact that there are constraints on a word's pronunciations. Furthermore, Kaplan's *Historical-Intentional Model* struggles with the communicative function of words, and makes it so that communicating agents might be, in general, unaware of what words are being uttered. Lastly, I'll argue that *Internal-Lexical Models* provide no notion of a public shareable word and leave some pronunciation variations unaccounted for.

The conclusion will be that that all of these views struggle in part for the same reason: they focus solely on the speaker, ignoring the role that the audience plays in determining which word an utterance is an utterance of.

3.2 Replication Models

Replication Models identify a word with a fixed phonological or orthographic form, which is replicated or instantiated by the word's utterances. The spoken word 'cat', for example, would be identified with the platonic phonological string /kæt/, and all of its utterances would instantiate this form. It is in virtue of matching or having the phonological form /kæt/ that utterances are said to be utterances of this word.

Since all utterances of the word 'cat' must match this phonological shape, all

utterances must also match each other in phonological shape. On the simplest and most common replication view, the word is the type of which its utterances are tokens, and two utterances are utterances of the same word if and only if they share or manifest the same a phonological form.

In principle, there is another kind of *Replication Model* on which a word is an abstract object that *has* a fixed phonological shape. On such a view, the word 'cat' need not be identified with a fixed abstract string of sounds, but it would be associated with one (just as I am not identical to my eye color, although I have one). But on both kinds of *Replication Models*, two utterances are utterances of the same word if and only if they match in phonological shape. These kinds of views are based in the idea that facts about how utterances *sound* determine which words they are utterances of. It is identity in pronunciation that unites a word's utterances.[1]

We find versions of these views in Quine [36, p.24] when he claims that 'Boston' has six letters,[2] in Susan Haack [22, p.75] when she claims, "[Expressions are] either a pattern which similar tokens exemplify, or a class of similar tokens", and in Brenden Gillon [21, p.162] when he writes, "An expression is a phonic or graphic form: the former has acoustic properties and the later spatial ones."

These views can explain nicely how agents come to recognize which word an utterance is an utterance of and when two utterances are utterances of the same word. This is because on *Replication Models*, all utterances of the word 'cat' will sound the same and manifest the same pronunciation. But a bit of reflection suggests that this kind of account immediately faces difficulty.

First, if the same word can be spoken and written, there will be little in

[1]See Donald Davidson [16, p.90].

[2]Kaplan reads Quine as meaning to claim that 'Boston' has six occurrences of letters, but only five letters [27].

common between the formal shapes of spoken and written utterances. Even if a word embodied a tuple consisting of a phonological shape and a spelling, (so that replicating either would suffice) tolerance among pronunciations or among spellings would be unaccounted for (e.g., that the same word is sometimes correctly pronounced [wɔtəˇ] and sometimes [wɔʔəˇ], or that the same word is sometimes correctly spelled *c-o-l-o-r* and *c-o-l-o-u-r*) [27].

It is worth noting that although spellings are in some sense symbolic representations of pronunciations, the connection between spelling and pronunciation is made complicated and unsystematic by the fact that spelling reform does not keep pace with sound change. Furthermore, there are many more variants in pronunciation than there are in spelling (e.g., there is typically one standard American English spelling, but many variants in pronunciation across the United States). Without some principle as to how far utterances and inscriptions can stray from each other, the view simply ignores the variation facts taken as data in this dissertation. As such, it is untenable for our purposes.

Second, utterances cannot be grouped solely on the basis of phonological form, since not all utterances that share a phonological shape are utterances of the same word (e.g., the words 'writing' and 'riding' share the pronunciation [raɪɾŋ], 'profit' and 'prophet' share the pronunciation [prɒfɪt], 'bank' and 'bank' share the pronunciation [bæŋk] etc.). Furthermore, the very same shape might tolerated as an articulation of one word in some context or dialect but not another. For example, an utterance [pɪn] made by an English speaker with a Southern American dialect and an utterance [pɛn] made by an English speaker without a Southern American dialect might count as utterances of the same word 'pen', while utterances of [pɪn] and [pɛn] made by two speakers without Southern American Dialects do not.

3.2.1 Phonemic Replication Models

One could attempt to develop a more sophisticated *Replication Model* by requiring that similarity in phonological form among utterances hold at a more abstract level. It might be that while utterances of the word 'pen' differ in their physical shape or sound properties, they all have the same form at what linguists call the *phonemic* or *underlying level*.

Modern phonological theory could lend some support to such a view. Traditionally, there are two levels of representation that are important in phonological theory: the phonemic or underlying level, which represents the basic sounds that a speaker associates with a word, and what we might call the phonetic or surface level, which represents the various ways these basic sounds will be articulated in context. For example, for English speakers, the sounds corresponding to the letter 't' in utterances of the words 'cat', 'stop', and 'little' all differ in their physical properties: in 'cat' the 't' is unaspirated, in 'stop' it is aspirated, and in 'little' it is voiced and flapped. But English speakers are often unaware of these differences and have the intuition that occurrences of these sounds are occurrences of the same sound. That is, they abstract away from some pronunciation variation and treat sounds that are physically different as the same. We may call this level of representation, at which phonetically different sounds are represented as the same sound, the *phonemic level* [24].

If all variation in pronunciation among a word's utterances were variation of the same phonemic form, then a replication model that held at the phonemic level might survive. For example, such a view might claim that although the utterances [pɪn] and [pɛn] have different acoustic properties, they are in fact replications of the same underlying form. (Although the triangle I draw and the triangle you draw might be different sizes or have different angles, we abstract away from these

differences and treat both of these inscriptions as inscriptions of the same basic shape).

Unfortunately, variation at the phonemic level is tolerated as well, and articulations of different phonemic forms are often grouped as utterances of the same word. For example, a Southern American English speaker's utterance of the word 'pen' (e.g., [pɪn]) and a Southern Californian speaker's utterance of word 'pen' (e.g., [pɛn]) are in fact thought to stem from different phonemic representations and are not simply the result of surface variation. There are many other examples, involving different kinds of phonological change, in which the number or distribution of phonemes or basic sounds is different for one group of language users than it is for another.

For example, in some parts of the United States, the name 'Mary', the verb 'marry', and the adjective 'merry' are all pronounced the same, but in some parts of Eastern Massachusetts, they are all pronounced differently. In some English dialects, the words 'cot' and 'caught' are pronounced differently and in some they are pronounced the same. These differences are thought to be the result of a difference in the distributions of phonemes, and not merely a result of the realization of the same phonemes in differing environments [24]. That is, these words may be abstractly represented differently by different speakers.

Since different phonemic forms can be associated with the same word, the problem of variation simply reappears at the phonemic level. We are left without an account of which phonemic variations are tolerable and which are not.

3.3 Intentional Models

Given that *Tolerance* prevents us from relating a word to its utterances via a match in phonological form, *Intentional Models* offer a radical alternative: that which

word is uttered depends on the speaker's intentions. After all, speaker-intention seems to be determinative in cases where there are two words pronounced in the same way (e.g., in determining whether an utterance [bæŋk] is an utterance of the word that refers to a financial institution or the word that refers to the land alongside a river). I'll focus on what I take to be the most well developed versions of these models—what I call *Historical-Intentional Models*.[3]

3.3.1 Historical-Intentional Models

Historical-Intentional Models reject the two fundamental claims of *Replication Models*: the first is the claim that a word is a kind of eternal platonic object, and the second the claim that a word's articulations are determined by phonological resemblance (or by any physical features of the utterance itself). Instead, *Historical-Intentional Models* place the origin and causal history of a words at the center of their individuation conditions and speaker-intention at the center of the pronunciation facts.

On David Kaplan's version, words are created objects with introduction events modeled after Kripkean baptisms for names [28, p.508], [27], [30]. All words have an initial stage, or first utterance, which is not intended to be an articulation of any previously articulated word [3]. These creation events mark the introductions of new words and provide the following individuation conditions: two utterances are utterances of the same word if and only if they are causally connected in the right way to the same first utterance [27, p.98], [28, p.509].

After a public word is created, it can be passed on to different agents via various communicative acts. For example, suppose agent A introduces the name W for her daughter. A can pass W on to agent B if (i) A makes an utterance

[3]Millikan [34] and Richard [38] also endorse historical models.

of W and (ii) B, upon hearing A's utterance, stores W in her mental dictionary (or lexicon). B stores W by introducing a mental linguistic structure, or lexical item, into her lexicon, and B's lexical item is then historically connected to W in virtue of the fact that it was introduced via certain historical interactions with utterances of W.[4]

According to *Historical-Intentional Models*, an utterance is an utterance of a public word W when the speaker, in making her utterance, intends to repeat W. There are different ways to flesh out what constitutes an intention to repeat W. On what I will call *the public view*, the relevant kinds of intentions are, as we might expect, straightforward intentions to repeat public words. On this version, a speaker's utterance is an utterance of public word W when she makes her utterance with the intention of repeating the public word uttered in a previous utterance of W. [23].

Kaplan gives the example of a subject who is asked to repeat the word uttered by a tester. He highlights the fact that if we believe the subject is earnest in intending to repeat the word, we will count his utterances as utterances of the same word the tester uttered. Even if the subject's utterances differ from the tester's in pronunciation, Kaplan says, we will say he is repeating the word *the best he can* [27, pp.102-105]. Accordingly, two agents' utterances will count as utterances of the public word W so long as (i) they both have lexical items historically connected to W and (ii) they both intend to repeat the public word W.

On what I will call *the basic view*, the intention to repeat a word W is fleshed out in terms of the intention to select a particular lexical item (which as it happens, is historically connected to W) [28, pp.514, 518]. On this view, the explanatory

[4]These interactions are not simply causal; they must also be of the right intentional kind [27, p.112].

intentions are basic intentions to select and perform operations over particular mental structures. We can conceive of basic intentions as commitments to perform concrete actions that (i) are over objects agents interact with directly and (ii) an agent can readily carry out.

Lepore and Stone characterize basic intentions as a kind of direct intention, whose objects must be specified/represented in indexical terms. Basic intentions are not necessarily conscious and are perhaps carried out by sub-personal modules [32, pp.215-218]. For example, intentions to perform certain motor skills, such as my intention to grab *that* thing in front of me or to walk to *that* building, constitute basic intentions, given the kind of direct control agents have over performing these skills.

Interactions with mental symbols and structures plausibly involve these kinds of intentions, since mental symbols are things that agents have direct indexical access to and agents can perform actions over them immediately. In the case of articulating a word, the relevant basic intention is the intention to retrieve or select *that* lexical item.

Accordingly, an utterance is an utterance of a public word W when its production involves the speaker's basic intention to select a lexical item historically connected to W in the manner described above. Additionally, two utterances are utterances of the same public word when the lexical items the speakers intend to select are historically connected.

3.3.2 Problems with Intention

If speaker-intention determines which public word has been successfully uttered, we can account for some variations predicted by *Tolerance*, including how there can be a single word pronounced [pɛn] by some agents and [pɪn] by others. On

the public view, this is because it is possible for two agents A and B to intend to repeat the same public word despite the fact that they articulate it differently. On *the basic view*, this is because it is possible that the lexical items A and B intend to select are historically connected, though articulated differently. Unfortunately, grouping articulations solely by speaker-intention generates pronunciation facts that are too permissive and violates *Constraint*.

On *the public view*, there are no constraints on how far utterances of the same word can vary in their phonological features [27, pp.101, 105].[5] Articulations [dɒg] and [kæt] could be counted as utterances of the same word so long as the speakers intend to repeat the same word. If the infamous Reverend Spooner intends to repeat the words 'The lord is a loving shepherd', he'll succeed in uttering them when he makes the utterance, "The lord is a shoving leopard" [28, p.520]. This is to claim that in some circumstances, the articulation [lɛpərd] is an utterance of the word 'shepherd'. Such a claim ignores the fact that there are standards with respect to how a word is articulated; it violates *Constraint*. Clearly, it is not enough for an agent to successfully utter a public word that she intends to do so.

Similarly, it's not enough for an agent to successfully utter a public word that she intends to select an appropriate lexical item. Suppose we allow that intending to select a lexical item is sufficient for selecting it. (After all, it seems plausible that, given the kind of direct access agents have to their lexical items and the kind of control they have over carrying out intentions to select one, very little can come between the intention to select a particular lexical item and successfully selecting it.) It is still controversial that if, for example, a speaker selects a lexical item historically connected to the public word 'dog', she'll succeed in uttering that word regardless of the sounds she makes. (Suppose she articulates her lexical

[5]Kaplan also makes exceptions for cases in which we would describe an agent as failing to make a speech act altogether [28, pp.518-519].

item as [kæt]). It's even controversial that she'll succeed in articulating that lexical item, regardless of the sounds she makes.[6] Again, to make such a claim ignores that there are standards with respect to how public words and lexical items are articulated.

We might worry that the kinds of cases I'm using to press against *Intentional Models* all involve slips of the tongue, or cases in which some part of the articulatory apparatus malfunctions. Consequently, an intentionalist might respond that he is only theorizing about cases in which things are not malfunctioning. But, we may consider a different kind of example in which the speaker's articulatory system is functioning properly, but the intentionalist view makes the wrong prediction.

Suppose that you introduce your daughter to me, and you tell me her name is 'Jess'. Suppose further that although I pick up this name from you and I intend to refer to your daughter by using that name, I think her name is 'Tess'. As a result, I refer to your daughter by uttering the sounds [tɛs]. On *the public view*, since I use my lexical item with the intention to repeat the word you've uttered, I'll count as having uttered your daughter's name. On the basic view, since my lexical item is historically connected to yours and I intend to select that very lexical item, I'll count as having uttered your daughter's name. This case involves no slip of the tongue or articulatory malfunction. However, I think you can rightly respond to my utterance of [tɛs] with, "That's not my daughter's name."

3.3.3 Problems with Historical Origin

I've argued that *Intentional Models* struggle with *Constraint* and the fact that there are standards with respect to how a particular word gets uttered and pro-

[6] At the very least, we will need to know more about how lexical items get articulated before accepting that any variation in pronunciation will be tolerated (more on this in 3.4).

nounced. Thus, adjudicating when two utterances belong to the same word by appeal to speaker-intention is not adequate. I'll end by briefly suggesting that the individuation of words in the communicative setting via their historical origin faces its own challenges.

First, it is questionable whether words generally have the precise introduction events that the view says they have. For example, at the present time, the English word 'cat' and the Latin word 'cattus' have a common history. It is also likely that the English word 'cat' never enjoyed its own autonomous baptism event, but instead evolved in a complicated fashion away from of its predecessors. Consequently, *Historical-Intentional Models* are in danger of blending all words that are cognates into one big word, resulting in far fewer words in English than the dictionary suggests and obliterating distinctions that matter for communication.

If we want to claim that English 'cat' and Latin 'cattus' are distinct words and that historical origin is sufficient for word-individuation, we need to find 'cat' it's own baptism event. But it isn't clear how to do this, given facts about how languages and words evolve from one another. This suggests that either (i) sharing an origin is not sufficient for word identity or (ii) word origins are not, in general, anything like what the view says they are like.

Kripke offers a possible solution when he highlights how the adoption of a new referring-intention can serve to create a new name. Suppose that I want to name my pet aardvark 'Napoleon' (after the French emperor). I am able to create a new name 'Napoleon' to refer to my pet aardvark so long as I do not intend to use it to refer to what previous agents intend their uses to refer to (e.g., the French emperor). When I break from the referring-intentions of those from whom I acquired the name and introduce a new referring-intention, I can effectively introduce a new name and start a new causal-intentional chain [30].

Perhaps Kaplan can appeal to a change in intention to give the English word 'cat' its own introduction event. The analogous story wouldn't appeal to a change in referring-intentions since 'cat' and 'cattus' have the same reference, but instead, it would require that speakers create a new introduction event by producing utterances with different *expression-intentions*. That is, somewhere in the causal chain, a speaker would have lost the intention to utter the same word that was uttered in a previous utterance. This break from the intentions of previous speakers might then mark the introduction of the new word 'cat'.

It seems plausible that currently, when I utter the word 'cat', I do not intend to utter the same word that was uttered by a Latin speaker's utterance of 'cattus'. But the closer we get back to the first uses of the English word, the more likely it is that speakers had a myriad of intentions, partly to go on in the same way as their predecessors, partly to speak as others do around them, partly to utter a certain phonological form, etc. If a historical-intentionalist wants to pursue this strategy, he'll need to fill out how exactly this is supposed to go.

Additionally, it is questionable whether distinctions in historical origin map onto the distinctions between words that speakers are sensitive to in a communicative exchange. Suppose two scientists A and B on opposite sides of the country independently discover the same species of tree, and both name it, 'belm'. In Kaplan's terms, the two words would be *phonographs*. A and B then pass on this name to various individuals, and after some time, two agents C and D, who have picked up this name from different causal chains, meet in conversation. Suppose C and D begin producing and interpreting utterances and come to communicate successfully. On *Historical-Intentional Models*, although C and D's utterances sound the same and mean the same, they belong to different words.

But suppose, what is likely, that when C and D begin communicating, they

start treating all utterances [bɛlm] as utterances of the same name. They might both come to adopt the intention to use the same name that the other is using and come to expect of each other that they are using the same name. In such a context, I think, it is controversial to claim that there are two distinct words in play despite the fact that communicating agents are not sensitive to this distinction. If the goal is to provide an account of words that meets the demands of philosophical accounts of communication, then a theory with this counterintuitive consequence is not adequate.

The scientist case further suggests that on *the public view*, the role of intention and the role of origin are in conflict with one another. On the one hand, C and D are set up by way of historical origin to utter different words, since their lexical items are linked to different baptism events. But on the other hand, once their conversation starts, they are set up by way of their mutual intentions to repeat/utter the same word as one another.

We might respond on behalf of Kaplan that before the conversation starts, C and D are uttering distinct words, but once it begins, they are uttering the same word. When C and D adopt the intentions to utter the same word as one another—when they begin to align their usage with each other—they might be engaging in exactly the kind of intention-change that can serve to introduce a new word. They might effectively break from the intentions that connected their articulations to different introduction events. By claiming that the agents have created a new word, Kaplan might have a way of claiming that the agents are using the same word.

While this maneuver will predict that C and D can use the same word, it mistakenly identifies a common origin as the underlying explanation. If a new historical chain may be created whenever agents align their intentions with one

another, then it seems that the word-individuation facts are more fundamentally grounded in this *other activity*—the adoption of mutual intentions and expectations to be using the same word. The fact that C and D are using a word with a common historical origin (as would be posited by this maneuverer) seems like it comes along for free, as a side effect of their adopting the right kinds of intentions. It's the intentions that seem to be doing the work, and while it might be true that a new word has been created, that fact seems to be explanatorily crowded out.

Lastly, couching word identity in historical origin makes a word's identity, in the majority of cases, opaque to its users. For example, consider the two scientists, who both baptize the same species of tree "belm", and thus according to the theory, will have made two distinct words with the same meanings and pronunciations. Recall that the two scientists were not in a position to know whether their own articulations were causally connected back to a single baptism event.

It is likely that most of the words in an agent's vocabulary are words whose actual historical-intentional facts are unknown to her. On the current account, it could turn out that successfully communicating agents are, in general, ignorant of whether two utterances are utterances of distinct words. But the possibility of this kind of pervasive ignorance is a detriment to successful communication. At the very least, it renders the natural picture of linguistic communication that goes via word-identification baffling.

The less controversial position is that (i) speakers in a conversation are not, in general, ignorant of whether two utterances in that conversation are utterances of the same word and (ii) the fact in (i) provides the possibility for agents to reliably content-match. These considerations suggest that the historical notion of a word will not suffice for our purposes.

Intentional Models are not without merit. They allow for pronunciation variation and provide an account of which homonym is uttered on a given occasion. Kaplan's version provides a thoughtful picture of words as public items capable of being uttered by different individuals over time. But while I take seriously the promotion of intentions (and in particular, intentions to be doing what others are doing) in accounting for the pronunciation facts, appealing solely to speaker-intention violates *Constraint*. Furthermore, positing public words only helps explain successful communication if speakers can reliably uncover which of these words is being uttered.

Historical-Intentional Models are natural in some contexts, for example, in giving diachronic accounts of words. It might be that Kaplan's model is necessary for trying to understand how words may change their phonological or semantic properties over time. But, with no constraints on how far utterances of the same word can vary and the possibility of ubiquitous ignorance with respect to the word-individuation facts, these models are insufficient for our purposes.

3.4 Internal-Lexical Models

Individual-Lexical Models offer a way to restrict *Tolerance*. They hold that principles of agents' internal grammars limit how far articulations of the same linguistic form can vary.

The notion of a word on *Individual-Lexical Models* is that of an entry in a speaker's mental dictionary. Following Chomsky, these entries, (also called *lexical items*) consist of certain phonological, syntactic, and semantic information [11]. For example, the lexical entry 'cat' might specify the phonological form /kæt/, the syntactic category *noun*, and a set of semantic features that describe cats. The list of a speaker's lexical items is called a *lexicon*, and it characterizes part of

an agent's *I-Language* [11], [9].

An I-Language is an individual's procedure for pairing linguistic expressions with meanings, and it embodies her ability to produce and understand an indefinite number of linguistic expressions. It is traditionally described as (i) intensional (it is a description of a procedure for producing and understanding linguistic structural descriptions), (ii) internal (this procedure is implemented in an agent's brain), and (iii) individual (this procedure, supervening on an individual's biology, is not shared by other speakers).

Part of an agent's I-Language consists of a phonological grammar, which can be thought of as an input-output device, taking as input the phonological component of a lexical item (call this the underlying form), and generating as output a representation of how that underlying phonological form will be articulated in context (call this the output form). Underlying forms can be thought of as representing the strings of basic sounds that are associated with a lexical item, and will be represented between slashes / /, and outputted forms will be represented between brackets [ ]

On this view, two articulations make by a given person belong to the same lexical item if the same lexical item taken as input to the grammar leads to those two articulations as output. I'll subsume these claims under what I will call *Internal-Lexical Models*.

3.4.1 Phonology and the Pronunciation Facts

Appealing to an agent's I-Language and in particular to the part of this procedure devoted to manipulating sounds, can predict rather nicely some of the pronunciation facts. This is because, while an I-Language is a capacity for generating and interpreting an indefinite number of expression-meaning pairs, it also respects

substantive constraints.

For example, while my I-Language pairs the interpretation that Mary loves Alice, but not that Alice loves Mary, with the form 'Mary loves Alice', it does not assign any interpretation to the string 'loves Mary Alice'. It pairs the string of sounds [dɒgz] but not [dɒgs] with the plural form of the word 'dog', and it pairs no interpretation with the string of sounds [bnɪk]. In this sense, I-Languages are procedures for generating and interpreting pairs of expressions and meanings in constrained but productive ways. To see how the present account can derive some of the variation and constraint facts, we turn to a brief discussion of modern phonological theory.

According to modern phonological theory, the articulations of words predictably vary in their speech sound properties in virtue of the fact that the articulations themselves are not atomistic units of sound. They are constituted by component speech sounds that get combined in various ways. This combinatorial sound system provides the possibility of encoding a large number of concepts into a smaller number of discriminable speech sounds [35].

When we attend to articulation variation at the level of these component sounds, the variation is often patterned and predictable. For example, in English, the 'd' sound becomes a tap ([ɾ]) when it is the single consonant between two vowels the second of which is unstressed (e.g., 'ladder' [lærr] and 'riding' [raɪrŋ]). Furthermore, these variations extend to new cases (e.g., we can predict that the plural in English of the novel common noun 'wug' is [wʌgz] and not [wʌgs]).

The study of the sound patterns in a language and how the articulation of a speech sound can vary systematically with context is called Phonology. It aims at understanding and describing the tacit system a language user uses in manipulating the sounds in her language. The Chomskyan perspective posits that this

phonological system is implemented in a particular agent's brain.

On the most prominent current phonology theory, Optimality Theory (OT), the fact that there is limited tolerance with respect to how far articulations of the same underlying form can vary is explained as the product of constraint interaction. OT posits that a phonological grammar is an input/output device involving (i) a function GEN for producing all the logically possible output forms for a given input, (ii) a function EVAL for selecting the optimal output as the actual output for a given input, and (iii) a lexicon, containing all of the underlying forms used as inputs to GEN [25].

Since GEN can generate any logically possible output for a given input, it is free to generate outputs [dɒg] and [kæt] for input /dɒg/. However, the second form will likely be ruled out as less than optimal by EVAL. EVAL is responsible for accounting for all observable regularities in sound patterning. It chooses the optimal output form by employing a set of ranked universal and violable constraints. Different grammars may rank the constraints in different ways, and lower ranked constraints can be violated only for the sake of satisfying higher ranked constraints. The grammar selects the output that commits the least costly violation of the constraints.

The constraints come in two flavors: markedness constraints and faithfulness constraints. Markedness constraints are universal output constraints that state marked or unmarked patterns, and they serve to preserve forms that are preferred by the perceptual and articulatory systems. For example, with respect to the feature [voice] for obstruents in coda position, [+voice] is unmarked and [-voice] is marked. This fact is reflected in the constraint **VOICED-CODA*, which states that obstruents must be voiced in the syllable-final position. Although avoided, marked structures are used, but only to create lexical contrast (to create forms

that provide the possibility of expressing different meanings). For example, although **VOICED-CODA* would predict that [and] is an unmarked structure and thus preferable to [ant], English makes use of both structures to express different meanings [25].

Faithfulness constraints require that outputs preserve the properties of the underlying form—that there be similarity between input and output forms with respect to some feature. Since these constraints limit input/output variation, they restrict how much articulations of a lexical item can vary, putting pressure towards a one to one mapping between meanings and forms. This is reflected in the constraint *IDENT-IO(voice)*, which states that output segments must match input segments with respect to voicing.

Internal-Lexical Models have the tools to explain some of the articulation variation facts rather nicely. First, since the set of phonological constraints used by a grammar can be ranked in different ways across different individuals, different grammars could choose different outputs for the same underlying form. For example, with respect to the underlying form /wɔter/, different constraint rankings could predict why one English speaker's pronunciation is [wɔrəˇ] while another's is [wɔtəˇ]). With respect to the underlying form /bɛd/, English grammars choose [bɛd] as the optimal output, violating the lower ranked markedness constraint (**VOICED-CODA*) for the sake of preserving the lexical contrast between lexical items /bɛd/ and /bɛt/. Dutch grammars choose [bɛt] as the optimal output, violating the faithfulness constraint for the sake of avoiding the marked structure [bɛd] [25, p.14].

Second, phonological grammars can explain how it is that the same lexical item can be articulated with different pronunciations depending on the linguistic context in which that lexical item occurs. The explanation for why the same input

form /bæd/ in some English dialects can be articulated as [bæd] in the utterance, "It's not so bad" but articulated as [bæg] in the utterance, "He's not a bad guy" is that different constraints transform the same underlying form depending on its linguistic environment. This particular transformation is known as *assimilation*: the word final 'd' in 'bad' assimilates to the place of articulation of the word initial 'g' in 'guy' [24].

Third, since a single individual's grammar limits how far an outputted form can stray from its underlying form, some articulation constraints can be explained by the principles governing constraint violation in an individual's grammar. For example, articulations [dɒg] and [kæt] won't be articulations of the same lexical item for any one speaker, since a single constraint hierarchy won't produce both as optimal outputs for a single input. Additionally, [dɒg] won't be an articulation of a lexical item with the underlying form /wɔter/ for any speaker, since it violates too many constraints without satisfying others. Specifically, it violates multiple faithfulness constraints without a decrease in markedness.

3.4.2 Problems with Internal-Lexical Models

Unfortunately, the claims and mechanisms posited by *Internal-Lexical Models* are inadequate for our purposes. According to the model, there is no sense in which two agents can articulate the same lexical item. *Internal-Lexical Models* operate within the assumptions of modern linguistic theory, which understands mental dictionaries and lexical items as *private* linguistic representations, posited as part of an individual's psychological architecture and implemented in an individual's brain.

There is no assumption that such representations are shared across individuals. Indeed such a picture is often resisted by claiming that nothing would be gained

in understanding aspects of a single individual's linguistic competence by positing that lexical items may be shared, or by positing that distinct agents may have the same lexical items. There is, according to this model, no notion of a common public word and articulations of lexical items are not in any sense articulations of shared structures.

Consequently, the Chomskyan approach to communication radically differs from content-matching models. Recall that in a communicative exchange, the fact that two agents both use the same words was supposed to explain how the audience could reliably come to entertain the intended content. On the current picture, however, there are no such public structures for both agents to use. Instead, each agent has their own I-Language, which may or may not have developed in similar ways. If both agents were exposed to similar linguistic input, then, given the initial biological similarities of their brains (what linguists call *Universal Grammar*), their grammars might be similar enough to facilitate communication. But a fleshed out story of how and when they will succeed is lacking and not forthcoming, as communication is not taken to be a central explanandum of linguistic theory in the generative tradition [11].

Part of the motivation for rejecting communication as an explanadum of linguistic theory comes from the idea that the study of communication is the study of how agents *use* their language for some social purpose. Theorizing about how agents do *this* is thought to be more properly about an agents' *linguistic performance* than it is about her linguistic competence.

But, it seems to me that there is an important difference between some of an agent's linguistic behavior when she is engaged in communication and other kinds of behavior that we might categorize as part of her linguistic performance. For example, the fact that I speak differently when speaking to a person of authority

than I do when speaking to a peer seems importantly different from the fact that I slur my speech when intoxicated. Indeed, the latter case seems like a performance *error* (as linguists often say) while the former does not. If anything, it seems as though I have a capacity for varying my speech in communication depending on whom I'm speaking to while I don't have a capacity for slurring while drunk.

I'm happy to concede that an explanation of an agent's linguistic behavior in communication might not be accounted for by the narrow understanding of an agent's linguistic competence that linguists are concerned with. But I'm not happy to concede that all of an agent's linguistic behavior in communication is merely the product of her performance, and not the product of an underlying capacity for speaking and understanding in a communicative context.

In fact, some of the work done by sociolinguists, such as William Labov, suggests that some of the linguistic behavior in communication that linguists ignore, or set aside as a product of an agent's performance, is highly structured when measured against certain social variables.

Some pronunciation variations are not accounted for by an individual's phonological grammar, and some of these variations are set aside by traditional phonological theory and categorized as free—which is to effectively abandon the project of trying to predict their distribution. Such variations are often tabled, not because they are unimportant, but because they are thought to be too erratic for the methods of theoretical linguistics [31, p.70]. For example, consider this quote from Labov describing previous observations of New Yorkers speech:

> "Indeed it was found that the speech of most individuals did not form a coherent and rational system, but was marked by numerous oscillations, contradictions and alternations which were inexplicable in terms of a single idiolect. For this reason, previous investigators had de-

> scribed large parts of the linguistic behavior of New Yorkers as being a product of pure chance, thoroughly haphazard" [31, p.124]

But, Labov continues, some of these seemingly unruly variations turn out not to be so when rigorously measured against the social and stylistic stratifications at work in the larger community:

> "But when the speech of any one person in any given context was charted against the overall pattern of social and stylistic variation of the community, his linguistic behavior was seen to be highly determined and highly structured."

For example, in studying articulations of words with consonantal 'r' in post vocalic-position (e.g., 'fourth', 'floor', etc.) made by employees from three socially stratified New York City department stores, Labov found that a speaker's use of 'r' correlated with which store she worked for: employees of higher ranked stores more often articulated 'r' (e.g., [flɔr]) than employees of lower ranked stores (e.g. [flɔ]). Labov found that such variation is predictable when mapped against a speaker's position along a social scale, and he predicts the more general pattern that if any two New Yorkers vary along some social scale, their differential use of 'r' will follow along the same scale [31, Ch. 2].

Labov also charted the employees' use of 'r' across both casual and emphatic speech styles, finding again that a robust pattern emerged: first, the rates of 'r' articulation increased with emphatic speech for employees of all stores, and second, employees of the lower-middle ranked stores *hypercorrected* their use of 'r' in formal speech. That is, while these employees had much lower rates of 'r' articulation than employees of the highest ranked store in casual speech, in formal speech, their rates of 'r' articulation were near or equal to those of the highest

ranked stores. Seemingly unpredictable variation can be understood in principled ways by appealing to social and stylistic pressures. Moreover, this data suggests—as will be a distinctive feature of my account—that the individual being spoken to is in some cases a crucial determinant of how a word is articulated.[7]

Or, to consider a more recent example, researchers found that in child-directed speech, the perceived gender of the child affected which pronunciation variations the adults used. It was found that when speaking to a female child, adults tended to use more standard pronunciation variants and those associated with prestige, while when speaking to a male child, adults tended to use less standard variants and pronunciations associated with vernacular speech [18].

Clearly, a speaker's I-Language influences the pronunciation facts and any theory ignoring this will struggle. But without a reliable mechanism for matching content, communication is unexplained, and without an appeal to social and stylistic stratifications, some variations are unexplained. Such variations, when measured against social factors, seem quite like the variations triggered by an agent's internal grammar—indeed, they are predictable and have consequences for communication. They also seem importantly different from other unruly variations, such as those resulting from speaking with a mouth full of food or speaking while intoxicated. In light of these considerations, maintaining that the social variations should be left unaccounted for is in need of support. I'll offer instead a theory that accommodates them, while also providing a way for agents to reliably match up their lexical items in service to their joint goal of communication.

[7]Likewise, Labov found that the centralization of the vowels [aʊ] and [aɪ] to [əʊ] and [əɪ] observed in articulations made by some inhabitants of Martha's Vineyard varied with the social attitudes of the speaker. Centralization occurred most in those who reject the speech style of those from the mainland and believe they impose on their way of life [31, Ch. 1].

3.5 Conclusion

None of the views exposited here adequately account for the desiderata outlined in Chapter 1. I'd like to suggest that we've been unsuccessful for the following two reasons: First, there is a view implicit in modern generative linguistics and philosophy of language that social factors and social variables are in some sense secondary. Facts about the social context, who our audience is, what our social attitudes are, and what our social stratifications are, are taken to affect our communicative exchanges, but in a pragmatic or post-linguistic way.

The general view is that there is a kind of hierarchy with respect to the tasks an agent must carry out in coming to understanding her interlocutor. This processing hierarchy generally requires that in coming to work out the intended meaning of an utterance, agents recognize which words were uttered first, and then attend to features of the social context and social variables after. The tasks an audience may goes through in recovering the intended meaning of an utterance might be ordered as follows: first, recover what sounds the speaker made, then recover what words the speaker said, then recover the linguistic meanings of those words, and then attend to the ways social variables can take you from the linguistic meaning to some other kind of pragmatic meaning.[8]

But I think this hierarchy is either false or misleading, and that we've failed to get a viable understanding of a public word because we've failed to see how the social variables can affect which words an agent has uttered. That is to say that the social context and social variables can matter at the level of what words a speaker has uttered, and that the audience must attend to them far earlier than the traditional view would require.

[8]See Clark [14] for an interesting version of this hierarchy in processing, and see Casasanto [7] for a discussion of how social variables influence language processing.

Second, in trying to understand the relation between a word and its utterances and when two utterances are utterances of the same word, the previous accounts appeal only to facts about the speakers who make those utterances and ignore facts about the audiences spoken to. They fail to accommodate (i) how features and expectations of our audience can temper how we speak and (ii) the role the audience plays in determining which word the speaker is uttering.

CHAPTER 4

The Communicative Policy Model

4.1 Introduction

In this chapter, I'll present an alternative framework that can accommodate both the variations described by theoretical phonology and those described by sociolinguistics. I won't start from the idea that a word is or has some fixed phonological form, which is then altered by various linguistic and social variables. I'll see the notion of a public word and the same-word relation as based in the idea that agents coordinate their private lexical items in the service to their joint goal of communication.

The animating idea behind my view is this: in trying to understand when two agents have uttered the same word, it doesn't matter whether those agents make the same sounds. Rather, it matters that they *interpret* each other's pronunciations in ways that pair up their internal lexical items. I'll argue that pronunciation variation poses no special challenge for this view, since lexical items may be coordinated despite being pronounced differently.

I'll aim to show that *The Communicative Policy Model*, which bases the pronunciation variation facts on pairs of overlapping communicative policies, respects *Tolerance* and *Constraint* while allowing that agents can use the same words and track their pronunciation variations. This framework significantly differs from previous models in providing an essential role for the audience in determining

which word a speaker has uttered.

4.2 Coordination

On my view, public words are based in sets of lexical items that are *coordinated*, in roughly David Lewis' sense. Lewis talks about a kind of problem agents face when (i) they must act together to achieve some mutual goal, (ii) there exist multiple strategies that would equally achieve this goal, but (iii) their interests would be best met by aligning their choices. Lewis calls these problems of interdependent decision *coordination problems* [33, p.24].

For example, drivers have a mutual interest in avoiding congestion and collision when on the road. To this end, they have multiple equally viable solutions: drive on the right (as Americans do), or drive on the left (as British do). It doesn't much matter which strategy they choose, so long as they align their choices. Crucially, their choice is arbitrary. There exist multiple solutions and mere rationality and world knowledge does not dictate that there is a "correct" solution.

Some of these coordination problems are what Lewis calls *signaling problems* [33, pp.122-133]. A two-sided signaling problem involves a communicator and an audience, where the communicator must decide on strategies for signaling and the audience must decide on strategies for acting on those signals. Lewis gives the famous case of The Sexton and Paul Revere needing to decide on contingency plans for communicating with one another about the plans of the Redcoats. They each face various options: to signal that the Redcoats are coming by sea, The Sexton could hang one lantern in the window or he could hang two, Paul Revere could react to seeing one lantern by warning the Redcoats are coming by sea, he could react this way to seeing two lanterns, etc.

What's important is that they don't care much which plan they choose, so long

as their plans combine to ensure that Paul Revere warns that the Redcoats are coming by land if and only if The Sexton observes they are coming by land, etc. So, each much choose a plan based on his expectations about what the other's plan will be.

When agents aim to use linguistic signals to exchange information, they face a coordination problem of this type. Linguistic communication can be thought of as a joint activity involving a speaker's strategies for articulating her lexical items and a receiver's strategies for construing articulations as her lexical items. But there are multiple strategies for articulation and construal that would serve the agents' joint goal of communication, and they don't care (in the long run) which strategies they choose, so long as they align their choices [33, p.50]. It might be too costly, for example, for an English speaker to adopt French overnight, but we would often prefer to adopt a new language in the long run than to give up on communication altogether. In this sense, the choice of strategies presents agents with a coordination problem. Each agent's choice of strategy is dependent on her expectations about the other's choice of strategy, and each would rather achieve their joint goal of communication than adopt a particular articulation or construal strategy.

Some sets of these articulation and construal strategies provide solutions to this coordination problem that are much like Lewisian conventions: they constitute a self-perpetuating system of preferences such that everyone has reason to conform so long as others do. Some sets might constitute solutions that are *coordination equilibria*, providing solutions on which no one would have been better off (with respect to the goal of communication) had any one agent alone acted otherwise [33, p.14].

For example, suppose that I have a strategy of articulating my lexical item x as

[pɛn] and you have a strategy of construing that articulation [pɛn] as your lexical item y, and additionally, you have a strategy of articulating y as [pɛn] and I have a strategy of construing your articulation [pɛn] as x. In such a case, no one would be better off with respect to our joint goal of communication had you, for example, adopted a strategy of construing my articulation of [pɛn] as y'. Indeed, we might have been worse off with respect to our mutual goal of reliable content matching, since y' might have a different meaning than y. Just as coordinated signaling strategies are essential for achieving The Sexton's and Paul Revere's shared goal of spreading the word about the Redcoats, coordinated communicative strategies are essential for achieving our joint goal of reliable communication. Importantly, each of us would rather meet this goal than adopt any particular articulation or construal strategy.

Since it is important to the conception of a coordination problem at work here that its solutions are arbitrary, it is worth pausing to defend the claim that sets of these articulation and construal strategies are arbitrary. An agent's strategies for articulating her private lexical items and for construing articulations as her private lexical items clearly flow from facts about her internal grammar and I-Language. I might pronounce a lexical item x as [wɔʔəˇ] because of a fact about the constraint hierarchy at work in my phonological grammar. But if this is true, how can an agent's strategies be arbitrary, in the sense required by Lewis' notion of a coordination problem? It may seem that I am not *free* to choose my strategies in the same way that The Sexton is free to choose his.

The first thing to appreciate is that I-Languages and phonological grammars themselves are arbitrary mechanisms. The implementation of a particular I-Language for an individual is not necessitated by genetics but develops as a result of linguistic input. Of course, for a particular speaker in a particular environment,

the development of her I-Language is not arbitrary. Given the initial state of her brain and her linguistic environment, she will adopt that particular I-Language. But, we might say that the implementation of any particular I-Language is arbitrary *for the species*, since it varies according to linguistic input.

The second thing to appreciate is that the communicative coordination problem is not entirely solved by individuals' I-Languages. As we saw in the previous chapter, some of an agent's strategies for articulation are the products of social variables, such as the prestige or gender of the person spoken to or the speaker's position along some social scale, and such strategies are not determined by her internal grammar. These strategies seem to be arbitrary in the traditional sense.

The question of how agents come to adopt sets of these strategies seems to me to be largely an empirical question, and I'd like to leave open that sets of these strategies could be established in a number of ways. I'll end this section however, by suggesting that agents might actively converge on them.

Whenever agents reliably solve a coordination problem according to a particular joint strategy (such as how to articulate private lexical items and how to construe articulations as private lexical items), they must be drawing on cues over and above rationality and world knowledge, which allow them to coordinate their behavior [32]. This is because world knowledge and rationality do not determine a unique solution.

Were agents to coordinate on sets of these strategies, this would require a mechanism that takes cues from at least the following sources: learned expectations about other agents, psychological mechanisms, and features of the social and conversational context.[1] This is because, as previously noted, the way agents articulate their lexical items and construe the articulations of others can vary along

[1] Lepore and Stone call this ability to draw on additional cues in coming to solve a coordination problem the exercise of a *social competence*.

all of these dimensions [32].

For example, A might adopt a particular construal strategy because of her expectation that B has a Southern dialect (recall the 'pen/pin' example), A might adopt a particular strategy because of the state of her internal grammar (recall the 'water' example), A might adopt a particular strategy because of the informal conversational context (recall Labov's studies), or A might adopt a particular strategy because of her position on some social scale (recall Labov's 'r' case). Were agents to have this kind of mechanism, we might see them as jointly converging on sets of strategies in coming to solve the coordination problem posed by linguistic communication.

I'll propose next that agents have exactly this kind of mechanism, a particular kind of competence for converging on sets of strategies as solutions to the communicative coordination problem.

4.3 Communicative Policies

I'll call a speaker's strategies for mapping her lexical items onto phonetic forms and a receiver's strategies for mapping received phonetic forms onto her private lexical items, an agent's *communicative policies*. I'll use the term *phonetic form* to refer to representations of how a linguistic form is to be pronounced, and we may think of phonetic forms as representations that interact with the articulatory and perceptual systems. Phonetic forms will continue to be represented using the standard International Phonetic Alphabet and they will be enclosed in square brackets. An articulation policy thus describes the sounds a speaker plans to make in sending a particular lexical item, and a construal policy describes the sounds a receiver plans to map onto a particular lexical item.[2]

[2]Phonetic forms do not describe the actual acoustic properties of the pressure waves a speaker produces when making an utterance. A description of these properties would be represented by

Communicative policies, as I'll conceive of them, describe a kind of Chomskyan *competence* [8]. They concern an agent's idealized capacities for sending and receiving phonetic forms in a possibly heterogeneous communicative context. This is to be distinguished from an agent's *performance*, which concerns her actual use of this capacity in context and might be affected by distractions, slips of the tongue, and other noisy features of the context. Communicative policies describe a kind of idealized competence for pairing phonetic forms with internal linguistic structures, and an agent may not have conscious knowledge of her policies. As such, her linguistic performance may deviate from her existing policies on any given occasion. We may think of the competence embodied in communicative policies as a *communicative competence*. This competence is to be distinguished from an agent's *linguistic competence*—the later being thought of as the proper functioning of an agent's I-Language. As we will see, communicative policies take account of social factors and the competence they describe extends beyond the functioning of an agent's internal grammar.

Importantly, communicative policies are sender, receiver, and conversational context specific. For example, speaker A might have a policy of articulating lexical item x with the phonetic form [hɪr] to agent B but a policy of articulating that same lexical item with the phonetic form [ɪr] to agent D. (Suppose B is a person of authority but D is not). Receiver E might have a policy of construing the phonetic form [pɪn] as lexical item y when sent by F but a policy of construing that same phonetic form as y' when sent by G. (Suppose F has a Southern American dialect and G does not). Additionally, speaker H might have a policy of articulating lexical item x with the phonetic form [flɔ] in a casual context but a policy of articulating that same lexical item with the phonetic form [flɔr] in a formal context.

what linguists call a *phonetic transcription*. My use of the term 'phonetic' concerns the abstract representations of sounds, and in this sense, involves a kind of idealization.

We can give the general schema for policies below (where S ranges over speakers, R over receivers, l over lexical items, p over phonetic forms, and C over kinds of social contexts):[3]

Articulation Policy: S has a policy of articulating l_s^i with p to R in C

Construal Policy: R has a policy of construing p as l_r^j when sent by S in C

Thus far, we may notice two ways in which the discussion has departed from our discussion of previous views: first, policies may be influenced by both features of an agent's phonological grammar as well as wider social features. The fact that A has a policy of articulating x as [pɪn] and not [pɛn] may flow from facts about her lexical items and phonological grammar, but the fact that some New Yorker's omit the 'r' sound in articulations of words such as 'floor' and 'fourth' may stem from a fact about their position along some social scale. Second, policies are what I will call *audience directed*, in that they vary with features and expectations of the audience.

Next, sets of overlapping policies can forge connective bonds between items in the private lexicons of distinct agents. We will say that $x \longrightarrow y$ for two lexical items x and y belonging to different agents A and B, when, in context C, A's policies of articulation combined with B's policies of construal connect x to y. For example, the overlapping policies in (i) establish that $x \longrightarrow y$:

(i) A has a policy of articulating lexical item x with phonetic form [wɔɾəˠ] to B in C

B has a policy of construing phonetic form [wɔɾəˠ] as lexical item y when sent by A in C

[3]Subscripts index lexical items to agents, and l_s^i indicates S's i-th lexical entry in her lexicon.

In (ii), $x \longrightarrow y$ and $y \longrightarrow x$

(ii) A has a policy of articulating x with phonetic form [pɛn] to B in C

B has a policy of construing phonetic form [pɛn] as y when sent by A in C

B has a policy of articulating y with phonetic form [pɪn] to A in C

A has a policy of construing phonetic form [pɪn] as x when sent by B in C

Notice that two lexical items can be connected via policy even though each might be associated via policy with a different articulatory phonetic form. In (ii) for example, x and y are connected although each is, as a matter of policy, articulated differently.

4.4 The Same-Word Relation

I'll now suggest that we may start to build up a notion of a sharable word as well as a notion of the same-word relation from sets of overlapping speaker and hearer policies. I'll divide up the core claims of *The Communicative Policy Model* into two theses. The first is what I will call *Public Word Constraint*.

***Public Word Constraint*:** Two lexical items x and y belonging to distinct agents A and B form the basis of a public word W only when:

(1) $x \longrightarrow y$ and $y \longrightarrow x$

(2) There is no lexical item x' in A's lexicon such that $x' \neq x$ and $x' \longrightarrow y$

(3) There is no lexical item y' in B's lexicon such that $y' \neq y$ and $y' \longrightarrow x$

The general claim here is that in order for there to be a public structure that two agents may use, those agents must have lexical items that are mapped onto one

another via their communicative strategies.[4]

The second thesis of the model is the *Same-Word Constraint*.

***Same-Word Constraint*:** Two utterances count as utterances of the same public word W only when they are policy adherent articulations of lexical items that form the basis of W.

In order for two utterances to count as utterances of the same word, first, they must be articulations of lexical items that meet (1)-(3) of *Public Word Constraint*, and second, they must conform to the articulation policies for those lexical items.

Notice that according to these constraints, while you and I don't need to have the same pronunciation policies in order to utter the same word, given that our policies are what they are, we do need to follow those policies. In this sense, sets of policies can establish what we might call *pronunciation standards*, which constrain which utterances and which pronunciations belong to the same word.

Furthermore, notice that no single policy establishes a pronunciation standard, but only sets of overlapping policies. It is not enough for an articulation to count as an articulation of a public word that it conforms to a speaker's articulation policies—it must also be mapped onto a corresponding lexical item in the receiver's lexicon by her policies. In understanding whether two utterances are utterances of the same word, it matters just as much how a receiver plans to construe those articulations as it does how a speaker plans to make them. The view advanced here thus not only appeals to the audience as a variable in a speaker's policies, it also appeals the audience's construal policies in accounting for facts about public words and their utterances.

[4] *Public Word Constraint* is based on Cumming's notion of *alignment* [15, p.9]. Where Cumming uses the notion of coordinated mental symbols to establish a notion of inter-subjective content, I'll use the notion of coordinated lexical items to establish a notion of inter-subjective syntax.

I've only given a necessary condition for two utterances to count as utterances of the same word, and I'll leave open in this dissertation what further condition might prove sufficient. And while the necessary condition I've given will face trouble, I'll argue that it is a good place to start.

In this paper, I've claimed that $x \longrightarrow y$ and $y \longrightarrow x$ (as opposed to simply $x \longrightarrow y$) are required to form the basis of a public word and thus for articulations of those items to count as articulations of the same word. This is because it seems to matter that y is connected back to x, and not, for example, to another lexical item in A's lexicon. Suppose A and B's policies are as specified in (ii) except that A is unaware of B's dialect. As a result, A has a policy of construing the phonetic form [pɪn] when sent by B in C as x' (suppose x' is a lexical item meaning *pin*) as opposed to mapping it onto her lexical item x (suppose x is a lexical item meaning *pen*). These policies are given below in (ii)$'$.

(ii)$'$ A has a policy of articulating x with phonetic form [pɛn] to B in C

B has a policy of construing phonetic form [pɛn] as y when sent by A in C

B has a policy of articulating y with phonetic form [pɪn] to A in C

A has a policy of construing phonetic form [pɪn] as x' when sent by B in C

In such a case, $x \longrightarrow y$, $y \longrightarrow x'$, but not $y \longrightarrow x$. The problem with this case is that x and x' could have (and in this case, do have) distinct meanings. We should be hesitant to say that articulations of x, x', and y are articulations of the same word, and in general, we should resist the claim that articulations of lexical items which may differ in meaning are utterances of the same word.

4.5 Examples

A few examples will serve to illustrate the animating ideas behind the current model and why we should want to require at least clause (1) of *Public Word Constraint*. Recall the policies in (ii), which establish that $x \longrightarrow y$ and $y \longrightarrow x$:

(ii) A has a policy of articulating x with phonetic form [pɛn] to B in C

B has a policy of construing phonetic form [pɛn] as y when sent by A in C

B has a policy of articulating y with phonetic form [pɪn] to A in C

A has a policy of construing phonetic form [pɪn] as x when sent by B in C

Furthermore, assume that these lexical items form the basis for public word W. As such, an articulation of x with the phonetic form [pɛn] made by A and sent to B in C conforms to A's policies connecting x to y, and thus, it counts as an articulation of W. However, an articulation of x with the phonetic form [pɪn] made by A and sent to B in C doesn't conform to A's policies and so, rather straightforwardly, it doesn't meet the articulation standards for W. Hence, it does not count as an articulation of public word W. Figure 4.1 maps out how the policies in (ii) connect A's and B's lexical items x and y.

Notice that an articulation with the phonetic form [pɪn] in C *does* count as an articulation of the public word W when it is made by B and sent to A, but not when it is made by A. This seems like a good result, since an articulation with the phonetic form [pɪn] made by A may be connected via policy to a distinct lexical item y in B's lexicon, as the Figure 4.2 depicts.

Next, recall an example from Chapter 2, in which my college roommate, meaning to say, "Let me interpret that for you!", utters the sounds, "Let me interpratate that for you!". According to *The Communicative Policy Model*, will her utterance

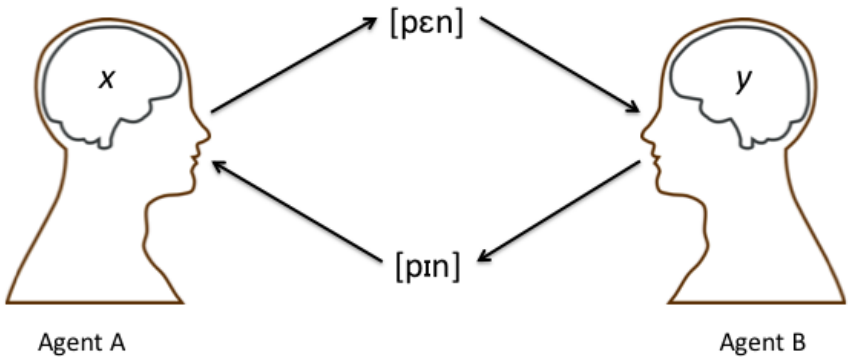

Figure 4.1: An example of $x \longrightarrow y$ and $y \longrightarrow x$

of [ɪntəprɛtat] count as an utterance of what I would call the public word 'interpret'? The most straightforward answer is no, and the reason is that it is unlikely that my roommate has a policy of articulating any lexical item as [ɪntəprɛtat]. After all, the case is supposed to represent a one-time malapropism or slip of the tongue.

It is thus likely that my college roommate and I have the policies specified in (iii) below, which map our lexical items x and y onto one another such that they may form the basis for the public word W.

(iii) Roommate has a policy of articulating x as [ɪntəprɛt] to me in C

I have a policy of construing [ɪntəprɛt] as y when sent by Roommate in C

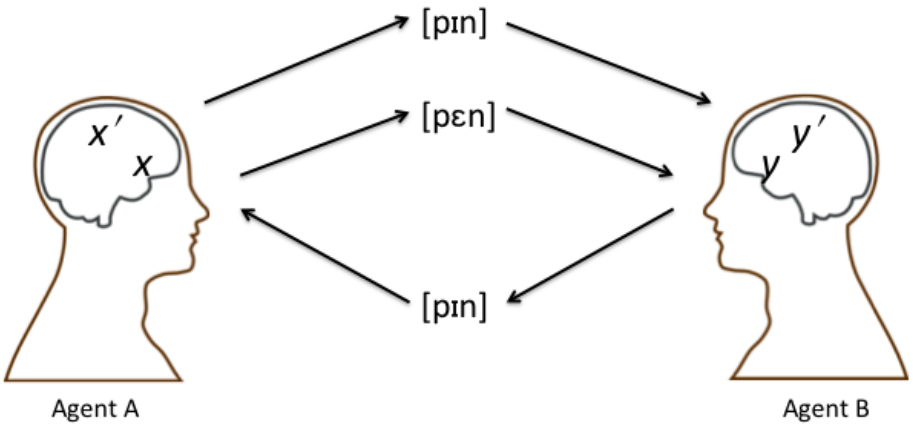

Figure 4.2: Another example of $x \longrightarrow y$ and $y \longrightarrow x$

I have a policy of articulating y as [mtəprɛt] to Roommate in C

Roommate has a policy of construing [mtəprɛt] as x when sent by me in C

Given these policies, my roommate's utterance won't count as an utterance of public word W, since her utterance deviated from our established policies.

Next, suppose Reverend Spooner and I have policies that map our lexical items x and y onto one another such that they form the basis for the public word W. The phonetic form [lʌvŋ] corresponds to what I would write as the word 'loving'.

(iv) Spooner has a policy of articulating x as [lʌvŋ] to me in C

I have a policy of construing [lʌvŋ] as y when sent by Spooner in C

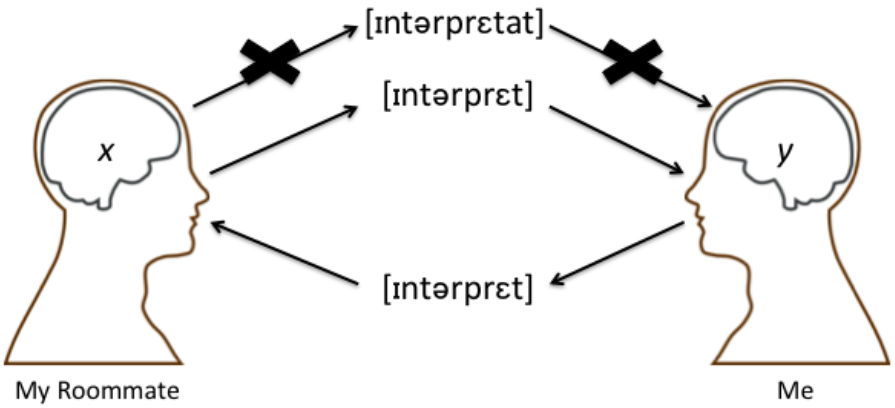

Figure 4.3: Communicative Policies for Malapropisms

I have a policy of articulating y as [lʌvŋ] to Spooner in C

Spooner has a policy of construing [lʌvŋ] as x when sent by me in C

Now suppose that Spooner, intending to say, "The lord is a loving shepherd" utters the sounds, "The lord is a shoving leopard." According to *The Communicative Policy View*, has Spooner uttered the public word W? Does his utterance of [ʃʌvŋ] ('shoving') count as an utterance of the same word as one of my utterances of [lʌvŋ] ('loving')? Figure 4.4 maps out how the policies in (iv) connect Spooner's and my lexical items x and y.

The most straightforward answer is no, it does not. And the reason why is because Spooner has deviated from his own policies, the policies which serve to coordinate our lexical items. His utterance doesn't meet the pronunciation

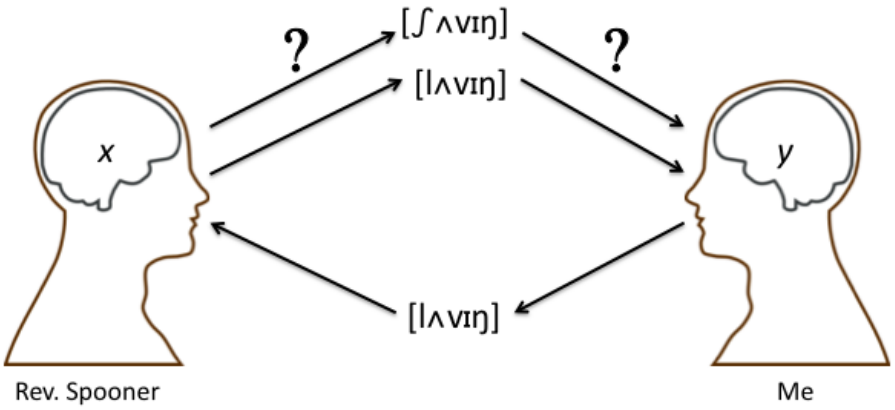

Figure 4.4: A Diagram of Reverend Spooner's Policies

standards set up by our policies.

But, we might say, surely I can understand Spooner. That is, I do interpret his utterance in the way he intended to be interpreted. So, why shouldn't we say that I do have a policy of construing an utterance with the phonetic form [ʃʌvŋ] when made by Spooner as my lexical item y?

This example serves to bring out another important point about the nature of communicative policies, as I'm conceiving of them. Policies aren't mere dispositions for articulation and construal. It might very well be that Spooner is disposed to habitually *mis-articulate* x, and I might be disposed to habitually repair Spooner's mistake in interpretation. Not all dispositions for articulation and construal serve to ground a communicative policy, but only those dispositions

carry certain kinds of future oriented commitments. [1]

For example, suppose that I have never encountered Spooner before, and I have no prior knowledge of how he speaks. I might still recognize that he meant to utter [lʌvŋ], and map the phonetic form onto my lexical item y in this instance, with no commitment to using this strategy in the future, should occasion arise. I might have no reason to think that this particular construal strategy will lead to successful coordination and communication in any future cases. In this case, I'll claim that I haven't established a policy of interpreting Spooner's utterances of [ʃʌvŋ] as lexical item y.

In this spirit, I'll say that an agent has established a communicative policy only if she is committed to returning to that strategy in future occasions, should they arise again. In the case in which I interpret Spooner by mapping [ʃʌvŋ] onto y, but I have no commitment to using the strategy in future communicative exchanges with Spooner, I don't adopt a policy of interpreting him in this way. And if there is no construal policy, if that phonetic form isn't mapped onto a corresponding lexical item in my lexicon, then we have reason to say that Spooner's utterance doesn't count as an utterance of the public word W.

Alternatively, suppose that Spooner and I are good friends and that we've spoken to one another numerous times. I might very well know that Spooner makes this mistake (and others like it) all the time, and so I might have reason to believe that this strategy will lead to successful coordination and communication in the future. In this case, I might establish a policy of construing Spooner's utterances [ʃʌvŋ] onto my lexical item y.

But notice that it takes more than my own construal policies to coordinate lexical items x and y. Successful coordination depends also on what Spooner's articulation policies are. What should we say about Spooner himself? Spooner is

known for making these kinds of pronunciation variations *all the time*. After all, they are called *spoonerisms*. In this sense, he is not like my college roommate in the previous example. Should we also say that Spooner has a policy of articulating x with the phonetic form [ʃʌvŋ]? If yes, then it may be that utterances of [ʃʌvŋ] and [lʌvŋ] are utterances of the same word, W. If no, *The Communicative Policy Model* will predict that Spooner has not made a policy adherent articulation of W.

I won't take a position here on whether Spooner as a policy of articulating x as [ʃʌvŋ]. Some things that might bear on our answer would be the nature of Spooner's commitments to articulating his lexical item x as [lʌvŋ], or to articulating x as he has in past communicative exchanges.

The most immediate trouble for the necessary conditions I've put forward is the existence of homonyms, or two words that sound the same. Clause (3) of *Public Word Constraint* requires that for two utterances to be utterances of the same public word W, they must be mapped onto and *only* onto each other by the agents' policies. But, homonyms make it such that an agent may have policies of articulating two distinct lexical items with same phonetic form or policies of construing the same phonetic form as two distinct lexical items.

For example, the verb 'profit' and the noun 'prophet' share the phonetic form [prɑfɪt]. This may mean that two agents do not have policies that align their lexical items in a one to one manner. Figure 4.5 represents how two agents' policies may lead to a many to many mapping between lexical items.

It isn't clear what we should do to accommodate these cases. We certainly want a view on which two agents may be using the word 'profit' in some of their communicative exchanges but 'prophet' in others. It would be undesirable to claim either that there is no public word in play or that there is one ambiguous

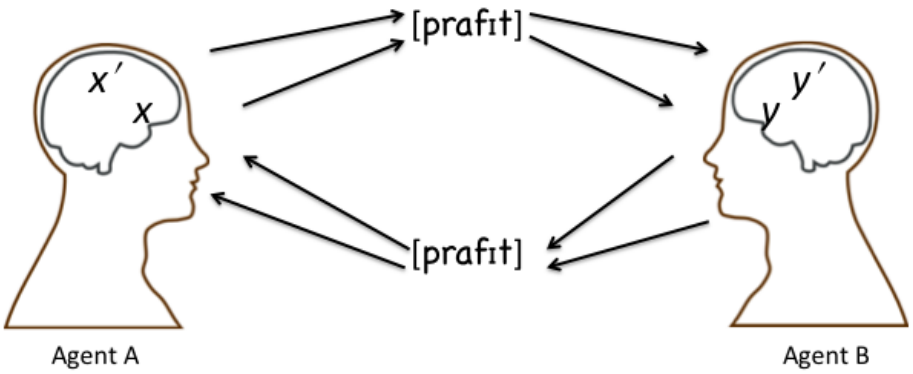

Figure 4.5: Communicative Policies for Homonyms

word in play when agents carry out these policies.

One option would be to relativize policies to a linguistic context as well as to a social context. We might amend our general schema so that agent A may have one policy of articulating lexical item x with the phonetic form [prɑfɪt] in linguistic context C (for example, when the utterance occupies the syntactic position of a verb) and a different policy of articulating lexical item x' with the phonetic form [prafɪt] in linguistic context C' (for example, when the articulation occupies the syntactic position of a noun). And similarly, agent B may have a policy of mapping the phonetic form [prɑfɪt] onto her lexical item y when it occurs in the verbal syntactic position, and a policy of mapping that same phonetic form onto another lexical item y' when occurring in the syntactic position of a noun. While

this solution may help with some homonyms, it won't help with homonyms of the same syntactic type (e.g., 'bank' and 'bank').

While requiring that lexical items meet conditions (1)-(3) in order to form the basis of a public word has these difficulties, it is, in general, a good place to start. First, we have some reason for wanting to guard against articulations of distinct lexical items belonging to a single agent as counting as articulations of the same public word. The worry is that if two of my lexical items, which are representationally distinct to me, could count as articulations of the same word, then I might be in general unable to tell when two of my articulations belong to the same word.

This is an undesirable result, when we consider that agents *are* able to tell when two of their articulations are articulations of the same word and when we consider the role words play in communication. We might think that there is for this reason a condition on the notion of a public syntax like the condition Frege imposes on content. Frege is said to have called for a notion of content that is *transparent*, in the sense that an agent is in a position to tell whether two of her concepts have the same content [19].

Second, we might want to require conditions as strong as (1)-(3) in *Public Word Constraint*, when we appreciate what happens when policies fail to meet (1)-(3). Suppose that for two agents A and B, their policies are such that B maps A's articulations of a single lexical item x onto two different lexical item y and y' (e.g., $x \longrightarrow y$ and $x \longrightarrow y'$). In such a case, A's and B's policies dictate that both y and y' are correct construals of x, but y and y' may have different meanings. In this case, we should be hesitant to say that policy adherent communication would be reliably successful, since arriving at a lexical item with the same meaning may be a matter of luck. On the contrary, when policies meet (1)-(3), there is exactly

one policy adherent construal of x and one policy adherent construal of y.

4.6 Meeting the Desiderata

Since an agent's communicative policies flow in part from her I-Language, the present framework can explain some variations described by *Tolerance* by appeal to an agent's phonological grammar. For example, the particular constraint hierarchy at work in an agent's I-Language can explain why she might have a policy of articulating x with the phonetic form [wɔɾəˠ] but not with [wɔʔəˠ]. Moreover, since agents' phonological grammars place substantial constraints on how far outputs of the same input can vary, they can place constraints on which policies hold for a particular agent. This can explain, for example, why A might have a policy of articulating x with [wɔɾəˠ] but not with [dɒg].

Additionally, while an agent's communicative policies flow in part from her I-Language, they are not wholly determined by it. Policies take account of external facts and can be changed or revised by the influence of external factors. Receiver B might have a policy of construing [pɪn] as y when produced by A but as y' when produced by D, or an agent might be aware that her interlocutor is a person of authority and vary her policies accordingly. We can also accommodate Labov's findings, by appealing to the fact that policies take account of the social and stylistic stratifications at work in the linguistic community. A might have a policy of articulating x with phonetic form [flɔ] ('floor) depending on social and stylistic pressures, or A might have a policy of articulating x' with the phonetic form [ɪr] ('here') depending on features of the conversational context.

On *The Communicative Policy Model*, the pronunciation facts can depend on a variety of linguistic, social, and stylistic factors, and can shift depending on the policies at work in a communicative exchange. As such, the model can

accommodate both the structured variations predicted by theoretical phonology as well as those predicted by sociolinguists.

While the framework allows corresponding lexical items to be associated with different phonetic forms, it requires that for articulations of those lexical items to belong to the same word, those forms must be mapped onto a corresponding lexical item by the receiver's policies of construal. In this way, the view further tempers *Tolerance* with *Constraint*. In (ii), for example, it is in part because A has a policy of mapping [pɪn] when articulated by B in C onto lexical item x that it counts as an articulation of the same word as A's articulations of x. Indeed, articulations with forms [pɪn] and [pɛn] might not count as articulations of the same word were A to have different construal policies.

Lastly, the competence embodied in communicative policies is sufficient to track when one public word has been articulated rather than another and when two articulations belong to the same public word. The tracking depends merely on the competent or correct functioning of agents' communicative policies. The model allows that agents can track the shifting articulation facts and avoids the consequence that communicating agents might be, in general, ignorant of which words are articulated.

I've advocated for an account of the pronunciation facts on which sets of overlapping articulation and construal policies provide a basis for a notion of a public word and set pronunciation standards. Such a framework is both *audience directed* and *audience involving*. It is audience directed in the sense that communicative policies take into account who is being spoken to and can shift depending on features of the audience. It is audience involving in the sense that whether two utterances are utterances of the same word depends not just on the speaker's articulation policies but also on the receiver's construal policies. Indeed, the basis

for a public word is given by sets of lexical items connected by overlapping speaker and receiver policies.

4.7 Public Words

It seems as though we have provided, or laid the foundation for providing, a solution to the philosophical problem about communication with which we began. We started with the claim that philosophical theories of communication appeal to the idea that two agents may use the same words and track which of those words are being used. I then argued that no previous model of words and the same-word relation met these desiderata while also accounting for pronunciation variation.

I then introduced a foundation for a notion of the same-word relation, arguing that agents may articulate the same word so long as they adhere to the communicative policies which serve to coordinate their internal lexical items. In some sense, I've claimed that communicative policies can provide a reliable mechanism by which agents may communicate.

Having done this, it may not be particularly important to choose a certain metaphysics of words. There may be further issues that settling on a particular metaphysics of words solve (such as questions about how quotation works, analyses of indirect discourse, etc.) but it seems that with respect to the challenge we set out to address, providing a reliable mechanism by which agents may communicate is sufficient.

I won't give an account of what public words are in this dissertation, but I will mention a few positions compatible with *The Communicative Policy Model*. One option is that public words are abstract objects related to their utterances via the overlapping communicative policies at work in a communicative exchange and whose existence is made possible by this network of communicative policies.

On such a view, public words would be the products of those very policies and based in sets of coordinated lexical items. Since the overlapping policies that give rise to the public word are solutions to a coordination problem, they might also give rise to a set of pronunciation conventions in roughly Lewis's sense, making the articulations that conform to policy into conventional pronunciation. This position, or something like it, is the position gestured at in 4.3 and 4.4, and it seems natural when we interpret *The Communicative Policy Model* as expanding *Internal-Lexical Models* by providing a mechanism by which agents can reliably align their lexical items.

A more radical position is that public words just are sets of overlapping communicative policies. Such a position might be natural if sets of communicative policies turn out to play the very role public words play on philosophical theories of communication. There are reasons in favor of exploring this view since sets of policies mediate between meanings and articulations, they solve a communicative coordination problem, and they are used by multiple agents. To be sure, this position would have some prima façe odd results since communicative policies are not the sorts of things we pronounce, write down, or look up in a dictionary.

Whatever metaphysics we choose, it does seem that the most natural picture of a public word on *The Communicative Policy Model* is one that sees a public word as more like a social contract or convention than it is like a fixed abstract form. This is because the conditions under which we have a shareable public word depend on the communicative principles of action that agents establish with one another. Just as we might say there is a social convention among a group of agents when there exists a jointly accepted principle as to how one is to act in a certain situation, we may say that there is a public word in use among a group of agents when there exist jointly accepted communicative policies for how one is to behave

in certain communicative situations.

Furthermore, just as a social convention might involve a complicated course of action in which agents will do different things (e.g., I will cook the food and you will wash the dishes) [20], the use of a public word might involve a course of action in which agents will do different things (e.g., I articulate my lexical item x as [pɛn] while you articulate you lexical item y as [pɪn], etc.)

4.8 Novel Words, Novel Pronunciations and Communication

Recall from Chapter 1 that the existence of novel pronunciations has been taken to threaten the role that conventionally established pronunciations play in explaining linguistic communication. Many traditional theories of linguistic communication take the fact that linguistic communication is reliable and efficient to implicate the existence of pronunciation conventions and agents' prior shared knowledge of these pronunciation conventions. However, if a speaker makes a pronunciation which is not yet conventionally associated with any word but manages to successfully communicate with her audience, it seems as though prior shared knowledge of a pronunciation convention played no essential role in explaining this success.

But a theory on which a word's pronunciations are determined by sets communicative policies could vindicate a role for pronunciation conventions in linguistic communication. Here's why.

First, consider that sets of communicative policies are in many ways like conventions: they are arbitrary, they solve a coordination problem, and they provide a self-perpetuating set of preferences that an agent has reason to conform to given that other agents also conform. As previously suggested, policies that coordinate lexical items belonging to distinct agents might not only provide a basis for a pub-

lic word, but they might also turn those pronunciations that conform to policy into conventional pronunciations.

Second, there is nothing preventing communicative policies from being *local*, *adaptive*, and established 'on the fly', as a conversation unfolds.[5] It is natural to think of communicative policies as holding between small groups of people in certain conversational contexts. Indeed, this has been the primary way of thinking about policies in this dissertation. The policies at work in a casual conversation between my sister and myself are likely different from the policies at work in an interview conversation between a potential employer and myself. The policies at work between speaker A (who has a New York English dialect) and speaker B (who has a Southern English dialect) are likely different from those at work between that same speaker A and another speaker D (who has a British English dialect).

There is nothing preventing policies from being adjusted in real time as a conversation unfolds, depending on various features of the audience, conversational context, and social context. I might realize part way through an exchange that I am speaking to the Dean of the university and adjust my articulation policies accordingly, I might pick up on your dialect after the conversation starts, adjusting my construal strategies accordingly, etc.

Just as agents might adjust their their communicative strategies as a conversation unfolds, they might adopt *new* policies altogether in light of learning new things about their audience and the conversational context. Imagine two speakers A and B who have never before spoken with each other. We may think of A and B as coming to the conversation with a set of policies that they each hope will facilitate successful communication. But just as those policies might need to be

[5]This view builds upon the work of Armstrong [1], in which a dynamic account of semantic conventions is developed.

adjusted, new ones might need to be added. A might need to add a policy of construing the phonetic form [pɪn] as her lexical item 'pen' when produced by B once she learns that B has a Southern English dialect.

Consider next what we might say about the related issue of accommodating novel words. Suppose that A says to B, "Can you mustard that toast for me?" Suppose further that B has never before heard or used the verb 'to mustard'. We may think of B as reacting to this utterance in the following way: she may first introduce a new lexical item, y, into her lexicon. She may then adopt the construal policy of mapping the phonetic form [mʌstərd] onto y when made by A. She may further adopt a new articulation policy of expressing y as [mʌstərd] to A.

Since there seems to be no real impediments to agents establishing new communicative policies 'on the fly', there is, in principle, no real impediment to agent's establishing novel words 'on the fly' (since on *The Communicative Policy Model* public words are based in sets of policies). But the issue isn't merely how agents establish new policies or new words, it is how they communicate via their use.

Recall that it is because you and I both have knowledge that the word 'Hesperus', for example, means *Hesperus* and that the pronunciation [hɛspərʌs] is a pronunciation of the word 'Hesperus' that our communication is successful in a reliable way. But before A's utterance of [mʌstərd], there doesn't seem to be any public word with the correct meaning for A and B to use (since before this utterance, there are no sets of overlapping policies involving the phonetic form [mʌstərd] which coordinate lexical items with the correct meanings and syntactic category.)

Likewise, before A's encounter with B's southern dialect and her utterance of [pɪn], there are no policies which coordinate A's and B's lexical items through

this phonetic form. But if prior to the relevant utterance, A and B either have no policies which coordinate their lexical items or only ones that would thwart communication, how can A and B communicate via those utterances (e.g., via B's utterance of [pɪm] or via A's utterance of [mʌstərd])?

Here are the seeds of an answer copacetic with *The Communicative Policy Model*. A's utterance of [mʌstərd] may simultaneously establish a new policy and propose a new public word. A may establish a new policy on the fly if she makes her utterance with the right future oriented commitments [1]. That is, she may adopt a new policy if she takes her utterance to be precedent setting—if she is committed to returning to this policy in future communications with B should they arise.

B may also adopt a new construal policy on the fly if she takes A's utterance to be precedent setting and adopts the right kinds of future oriented commitments. But until B does her part, there is no sense in which A, having uttered [mʌstərd], has uttered a public shareable word. A and B have to make the public word together. They have to coordinate their policies. In this sense, while A may establish an articulation policy on her own, she cannot unilaterally establish a new word on her own. If B proceeds by introducing a new lexical item into her lexicon and a new construal policy that coordinates this lexical item with A's, then they will have made the basis for a new public word. And if B does this, she may then use those communicative policies to interpret A's utterance, and communication will succeed.

But B might not proceed in this way. She may react to A's utterance of [mʌstərd], for example, by saying, "What? What do you want me to do?", or she might reply, "That's not a word". Or, she might introduce policies that do not successfully coordinate A's and her lexical items. (Suppose she thought A

was uttering the word 'butter' but in a new way, and so she adopted the strategy of mapping the phonetic form [mʌstərd] onto her lexical item 'butter'. Such a strategy might not result in coordination.) In this sense, it is only when the audience does her part that the speaker and hearer may both act by first making a new word and then proceed by using it in communication.

While further work still needs to be done on exactly how agents set up policies on the fly, *The Communicative Policy Model* seems naturally fit handle these cases. It already conceives of agents as working together to make the basis of a public word, and it conceives of communicative policies as flexible, dynamic strategies. If there is no real impediment to thinking of sets of communicative policies as conventions, then *The Communicative Policy Model* may maintain that pronunciation conventions play an important role in explaining communication, but also that agents may set up and use new ones 'on the fly'.

4.9 Conclusion

The Communicative Policy Model provides an account of the same-word relation that respects *Tolerance*, *Constraint*, and the manifest communicative function of language. It offers new insight into the nature of public words as based in the communicative policies of both speakers and receivers. The model succeeds where other models do not in part because it assigns a central role to the audience and in part because policies take account of both internal linguistic facts and external social facts. I've advocated for such a view by arguing that sets of communicative policies can effectively determine pronunciation standards, connect lexical items in the lexicons of distinct agents, and solve a communicative coordination problem.

The picture of words advanced by this chapter is not subject to Chomsky's criticisms. Recall that Chomsky objected to the notion of a public shareable word

invoked by philosophers, claiming that (i) there is no principally linguistic way to divide up the linguistic communities that are said to use them and (ii) this notion is divorced from the subject matter of linguistic science.

But *The Communicative Policy Model* doesn't divide up linguistic communities said to be using the same word on merely geo-political or arbitrary grounds. It appeals to the overlapping communicative intentions and strategies of groups of linguistic agents, claiming that a group of agents can be said to be using the same word only when their communicative policies have a certain profile.

Furthermore, the account offered here is not divorced from linguistic science. It sees public words and the same-word relation as based in sets of coordinated internal linguistic representations, which are posited by linguistic science, and it sees the coordination of these representations as influenced and constrained by individual phonological grammars, which are also posited by linguistic science. To be sure, the view appeals to capacities beyond those posited of linguistic science, but it see these wider capacities as intimately tied with the linguistic ones.

The Communicative Policy Model asks that we understand the relation between public words and their articulations by appealing to the interconnected communicative policies holding between agents in a linguistic community. Once we do this, we see that the articulation facts are fundamentally social facts—facts that cannot be understood by inspecting speakers' intentions or internalized grammars in isolation from the communicative contexts in which they operate.

There is a way of interpreting the model proposed here on which any time two agents speak with one another for the first time, they must *make* the public words they seek to use. They must establish sets of policies that coordinate their lexical items, providing the basis for a common public word. Until these policies are established, there is perhaps no sense in which both agents are using the same

word, and it might be that many more communicative exchanges turn out to involve word-making than we had previously thought.

CHAPTER 5

Questions and Replies

5.1 Introduction

Thesis 1: Private Words

A word is an internal linguistic representation playing a certain functional role in an agent's idiolect (or an individual's way of speaking and understanding). Each individual has her own mental dictionary, in which she stores the arbitrary associations between pronunciations and meanings in her idiolect. These entries are her private words. Entries in an agent's mental dictionary may or may not be similar to the entries of other agents in her linguistic community.

Thesis 2: Public Words

A word is a public object used by a group of agents for the purposes of communication. Groups of agents can come to successfully communicate to the extent that they use the same words with the same meanings and are able to recognize which words are uttered. These words are shareable, can be used by more than a single agent, and do not exist in the minds of the agents who use them.

Question:

What do private words have to do with public words?

Answer 1:

> These notions have nothing to do with one another. They are two different notions of a word aimed at explaining different phenomena. The notion offered in Thesis 1 is posited to explain certain aspects of an agent's linguistic competence, while the notion offered in Thesis 2 is posited to explain how agents come to understand one another via the use of linguistic signals.

Answer 2:

> Throughout this dissertation, I have offered a different answer: that public words are grounded in coordinated private words. I've argued that agents may coordinate their private words by converging on communicative policies in service to solving a coordination problem posed by linguistic communication. According to this view, groups of agents may use the same public word in virtue of aligning their private words.

In this Chapter, I'll survey some questions about the view offered in Answer 2 and reply to each in turn.

5.2 Questions and Replies

Question (1):

You say that whether my utterance 'cat' and your utterance 'cat' are utterances of the same word depends in part on the whether they are utterances of private words which are coordinated. But doesn't the explanatory priority go the other way around? What makes it the case that you and I both utter the public word 'cat' is that we have acquired, through exposure to various linguistic input, private analogues of the same English word. It is intuitive that we posit the public word first and then understand our private words in terms of which public words

they are instantiations of. You're view has it backwards.

Reply:

While the view you offer may be intuitive, it faces difficulties. According to the view you offer, in virtue of what is my lexical item an instantiation of the public English word 'cat'? What is this relation between public and private words grounded in? One option is to say that my private word is an instantiation of the English public word 'cat' because it is pronounced [kæt] and means *cat*. But this cannot be right, since many private words will count as instantiations of the English word 'cat' but have different pronunciations.

Perhaps it is that my private word is an instantiation of the English public word in virtue of the fact that I acquired my word from others who themselves have private English analogues. On this view, the relevant relation is a kind of causal one. But this cannot be right either, once we consider how languages evolve from one another. At some point this causal chain goes back to agents with non-English private words. There seems to be no good way of understanding what makes the private words instantiations of the public words by positing the public word first, and hence no good way of answering what makes two utterances utterance of the same public word.

My alternative proposal—that we understand the public word in terms of a kind of social relation that holds among private words—does not face the same difficulties.

Question (2):

According to your view, I can only use the same word as another agent if there exist communicative policies between us that coordinate our lexical items.

It seems that I have no policies with Groucho Marx. If this is correct, does it mean that I cannot use the same words he used? Doesn't it seem correct to say that I may use the same words as people with whom I've never directly communicated nor ever will?

Reply:

First, it seems correct to say that you *do* have policies that allow you to interpret Groucho Marx. It may be that you have, in addition to your specific policies for individuals with whom you've communicated directly, more general policies. These general policies (say, for speaking to and interpreting English speakers) might be what you use to interpret Groucho Marx, while you might use a set of more specialized policies when communicating with your neighbor.

Second, I'll also allow lexical items to be coordinated not only by direct communicative policies, but also by *indirect communicative paths* through the wider community. This allows that the lexical items of agent A and agent C might be coordinated even if there exist no direct communicative policies between them.

Following Cumming [15], I'll allow that two lexical items x and z may be connected by a *communicative path* when: there exists a sequence of lexical items beginning with x and ending with z such that each lexical item in the sequence is connected by direct communicative policies to its successor. For example, a lexical item x in A's lexicon could be connected by an indirect path to lexical item z in C's lexicon via another agent B (with lexical item y) if $x \longrightarrow y$ and $y \longrightarrow z$. We may allow that x and z are coordinated if either they are connected by direct policies or if they are connected by indirect paths.

In this sense, although you may not have any direct communicative policies with Groucho Marx, the state of the communication network at a given time might

be such that your lexical items are coordinated with his via these indirect paths.

Question (3):

You criticize *Historical-Intentional Models* for being in danger of predicting that there are far less words than a dictionary would claim. But your view would carve up public words in non-traditional ways as well. Aren't you in danger of predicting that there far are too *many* words? Where we might want to say there is the same English word being used across various linguistic communities, you might say there are distinct words for each community.

Reply:

First, while I've suggested that we look to sets of policies as a way to individuate words, I've only extensively argued for a necessary condition on two agents using the same word. But it may be right that *The Communicative Policy Model* makes more distinctions than our pre-theoretical intuitions would make.

For instance, suppose agents A and B have policies coordinating their lexical items x and y, agents C and D have policies coordinating their lexical items w and z, but A and B are isolated from C and D (that is, A and B have no policies connecting their lexical items with those of C or D and vice versa). Suppose further that each of the lexical items x, y, w, and z all have the same meaning and the same pronunciation (they are all articulated with the same phonetic form). We might be inclined to say here that all of these agents are using the same (e.g., English or French) word, but yet, *The Communicative Policy Model* might say they are not.

I think *The Communicative Policy Model* has the resources to say one of two things, depending on how the model is fleshed out. The first thing this kind of

view could say is that there may be indirect communicative paths (e.g., through a fifth agent E) that serve to coordinate A's and B's lexical items with those of C and D. In this case, we may have the resources to say that all the agents are in fact using the same word.

But it may be that there are no paths, direct or indirect, coordinating all of the agents' lexical items. In this case, I can imagine a proponent of *The Communicative Policy Model* claiming these agents are not all using the same word. And I can imagine a proponent arguing that this result is correct. Doesn't it seem correct that Martians, who have had no contact with Earthlings, cannot speak the same language as Earthlings? At the very least, historical-intentionalists will agree. Where historical-intentionalists will explain this claim by appealing to a lack of historical connection, the policy theorist will explain it by appealing to a lack of coordination.

However, I can also imagine a policy theorist resisting this claim, arguing that it very well may be the case that Martians and Earthlings can use the same words despite their physical isolation. We might take the position that lexical items can be coordinated without agents engaging in the activity of coordinating with one another. Here's what we might say.

Imagine taking a synchronic snapshot of A's, B's, C's, and D's policies. Some of these policies might be agent specific (some might be only policies for communication with B, for example), but some of them might be general (some might be policies for communication with any one in the United States, for example). It may be the case that while A's and B's agent specific policies coordinate their lexical items and create what we might call their own micro-language, their general policies also coordinate their lexical items with C's and D's. Whether the policy theorist wants to argue this depends on whether she thinks there are both

agent specific policies as well as more general policies, but I think the distinction is intuitive. And it might be the case that our general policies serve to coordinate our lexical items in ways that would individuate words in line with how words are grouped as *English*, or *French*, or *Mandarin*, for example.

Question (4):

Why can't we reduce the notions of a public word and the same-word relation to the notion of overlapping lexical items? That is, why shouldn't we say that a public word is an abstraction from lexical items that are sufficiently alike, and that two agents are using the same public word when their lexical items are sufficiently alike?

Reply:

There are two problems with this kind of view. First, phonological overlap is not sufficient, since lexical items may share a phonological form but different meanings (recall the case of homonyms 'profit' and 'prophet'). Second phonological overlap is not necessary since there can be wide differences in pronunciation of the same word. I've argued extensively in this dissertation that looking to phonological overlap struggles to account for pronunciation variation. Agents' lexical entries may differ in their phonological properties, and while I may have an entry pairing the phonological form [pɛn] with a certain meaning, you may have no such entry. In fact, you may have no entries involving the phonological form [pɛn] at all.

Second, you and I could have overlapping lexical entries by *accident*. And, while communication might succeed in cases of overlap, it wouldn't be reliable in the way that linguistic communication is. Recall that this was part of why

Internal-Lexical Models were unsatisfying. There was no reliable mechanism for establishing overlapping lexicons. On the view offered in this dissertation, it isn't the mere overlap of lexical items that matters, but the coordination of these lexical items achieved by the existence of communicative policies. On such an account, communication is not accidental in the way it might be on Chomskyan accounts.

Question (5):

There are other kinds of cases (aside from the Groucho Marx case) in which two agents might fail to have direct policies that coordinate their lexical items. Suppose my handwriting is so terrible that you cannot understand it, or suppose that you cannot understand my dialect, so much so that we cannot set up any policies at all. Does this mean that neither of us has used any public words in trying to communicate with one another? It is odd to claim that I have failed to say or use any public words simply because you fail to understand me.

Reply:

For some of these cases, appealing to indirect paths connecting with the wider community might give us recourse to claim that a public word is being used despite the fact the direct communicative policies fail to coordinate our lexical items. For example, even though I cannot understand your dialect and consequently don't have any construal policies connecting your articulations to my lexical items, our lexical items might be coordinated via indirect paths through the wider community. In virtue of this, you might still be using a public word.

But where this retreat to the wider community isn't possible (suppose there is no wider community or suppose that other communicative paths still don't establish coordination), a policy theorist may conclude there is no public word in

play.

Question (6):

Recall the infamous Reverend Spooner who intends to utter the words 'The lord is a loving shepherd', but makes the utterance, "The lord is a shoving leopard". Which words has Spooner articulated on your view? Since he hasn't made a policy adherent articulation of his lexical item 'loving', you won't conclude that he articulated the public word this lexical item forms the basis of. But it seems correct to say that Spooner accidentally uttered 'shoving'. How is this possible on your view?

Reply:

Notice that the phonetic form Spooner articulated (e.g., [ʃʌvŋ]) is connected via policy to a different lexical item, 'shoving'. Whether we want to say that Spooner uttered the public word 'shoving' might depend on the nature of his mistake. Suppose the situation is such that Spooner suffered an embarrassing slip of the tongue. In this case, it is helpful to make a subtle distinction between (i) what word has been *competently articulated* and (ii) what word is *represented* by a particular articulation.

Consider the following analogy: suppose we come across the inscription *OAR* on a piece of paper. Suppose further that the typist meant to type the letters *PAR*, but her fingers slipped, resulting in a typo. The question of what word is represented by this articulation is perhaps different from the question of what word this articulation is a competent articulation of. It might be that the word 'oar' is represented by the markings on the paper but also that those markings don't correspond to a competent articulation of any word. It seems natural to say

that the typist neither competently articulated the word 'par' nor the word 'oar'. After all, her fingers slipped!

In the same spirit, we might say that the word 'shoving' is represented by Spooner's utterance [ʃʌvŋ], while his utterance is not a policy adherent utterance of this word (nor of any other word). So while my view might deny that Spooner made a policy adherent articulation of his lexical item 'shoving', there is a sense in which the utterance he made, were it competent, would have been an utterance of the word 'shoving'.

Perhaps we laugh when Spooner speaks because we recognize that his utterances represent words that, if assigned their standard meanings, would mean something ridiculous, but we hesitate to assign his utterances those meanings because we hesitate to say that he competently articulated those words.

Question (7):

And what will you say about Yogi Berra when he utters "Texas has a lot of electrical votes" and we understand him? Will you adopt a solution in the Davidsonian spirit on which we establish a new convention (or passing theory) that the word 'electrical' means *electoral* (in this instance)? Or will you adopt a solution in the Kaplanian spirit on which we establish a new convention that the pronunciation [ʌlɛktrɪkəl] is a pronunciation of the word 'electoral' (in this instance)?

Reply:

It seems that *The Communicative Policy* could adopt either of these solutions, so long as we accept that agents can adjust and adopt new policies on the fly. However, we don't have to adopt either. We can adopt a less controversial position which has its seeds in our answer to the previous question. We may say that Yogi

did not competently articulate any word by his utterance [ʌlɛktrɪkəl]. We do not introduce any new policies or conventions (unless we discover that this is Yogi's way of pronouncing the word 'electoral' and not a mere slip of the tongue.)

Instead, we may say that communication succeeds here by some other mechanism. We recognize that Yogi's utterance contains a mistaken occurrence of the word 'electrical', we silently repair this mistake, and then assign the repaired sentence its conventional meaning. Yogi's utterance is like my roommate's, not like *A*'s utterance of, "Can you mustard that sandwich for me". I think there is something to be gained by keeping these cases separate, and by not treating all cases of novel variation in the same way. Encountering Yogi Berra for the first time seems importantly different from encountering the verb 'to mustard' for the first time. In the later case, but not the former, it seems that a canonical precedent is created that one has reason to conform to in the future.

Question (8):

We often say that someone is pronouncing a word incorrectly, or that someone made a mistake in pronunciation. How do you account for this? On your account, what sense is there to the idea that someone could being pronouncing a word incorrectly? Aren't my policies simply whatever my policies are?

Reply:

On my account, there is a sense in which an agent may make a mistake when she deviates from her own policies in making an articulation. And perhaps there is a sense in which agents ought not deviate from their established policies. Since these policies solve a coordination problem, and since others expect one to adhere to them, agents have a reason to adhere. But there is no sense in which someone

may have mistaken or incorrect policies.

We might say that there are better or worse policies given agents' joint goal of communication, since some policies will establish coordination and some will not. And there are perhaps prudential reasons for adopting policies that make coordination easier to achieve. But if you and I have coordinated our lexical items via our policies, there is no sense in which these policies ought to be different or that one of us has incorrect policies.

In this sense, whether a pronunciation is a mistake depends on the state of the communicative network and whether that pronunciation is licensed by established policy.

Question (9):

Suppose we have two agents A and B who have coordinated their lexical items x and y and made the basis for a public word W. Suppose now a third agent C joins their conversation and x and y become coordinated with C's lexical item z. Do we have the basis for a new word W' in this case? When do we form the basis for a new word as opposed to expanding the basis of an old word? Don't we need some notion of historical connection to track the identity of a word over time while allowing the lexical items of the coordinated set to change and evolve?

Reply:

This worry akin to a worry about when we would say people are making a new convention and when we would say they are following an old one. We don't think the convention to drive on the right changes into a new convention when more people begin driving on the right, even though the convention is, in a sense, based in the desires and preferences of the individuals. There is some way of

thinking about conventions such that they don't lose their identity when more or less people come to act in accordance with them.

I'd like to be able to claim the same intuitive point here, although I haven't said anything about how to do it, since I haven't said anything about how to track the identity of a public word over time. The theory presented in this dissertation concerns the synchronic state of a communication network and when we may say a public word is in use given a stable communicative network. Connecting networks over time is an interesting but different issue.

Question (10):

What does your view have to say about Kripke's famous Peter, who, overhearing talk about Ignacy Jan Paderewski on two separate occasions, comes to think that there are two Paderewskis, one who is an accomplished musician and one who is a politician [29]. On traditional historical views, since there is one man who has been baptized once, there is one public name 'Paderewski'. Peter has made a mistake by introducing two distinct lexical items for this single public name. What will your view say here?

Reply:

Historical models may be able to explain why it is that Paderewski is not irrational when he utters, "Paderewski is a musician and Paderewski is not a musician". Since Peter was unaware that the utterances he heard were in fact connected to the same baptism event, he has two representationally distinct lexical items which he is unaware are stages of the same public word. But notice that this model claims that that two of Peter's representationally distinct lexical items belong to or are part of the same public word, and that Peter is unaware of this.

I've already argued that this is an undesirable result when considering the role of public words in enabling reliable communication.

The pressing issue for us is not whether Peter utters a contradiction or whether he is irrational, but what Peter's policies are and whether he can communicate with other agents via those policies. For example, suppose Peter and a non-confused agent, B, try to communicate via utterances of the sounds "Paderewski". What happens when Peter utters, "Paderewski is a musician" to an agent who only has a single lexical item naming Paderewski? What happens when the non-confused agent B utters, "Paderewski is coming to town today"? Does communication succeed? If so, how? And if not, why not?

This seems to a case in which Peter's and B's communicative policies do not coordinate the relevant lexical items. Figure 5.1 represents what the communication network involving B and Peter might look like with respect to B's single lexical item for Paderewski and Peter's two lexical items for Paderewski.

Notice that there is no single lexical item in Peter's lexicon such that it is mapped onto and only onto y and y is mapped back onto and only onto it. There is also no item in A's lexicon such that it is mapped onto and only onto some single item in Peter's lexicon which is then in turn mapped onto and only onto it. That is, B's and Peter's lexical items do not meet clause (3) of *Public Word Constraint*, and don't form the basis for the same public word. Perhaps it is this feature of their communication network that rules out B and Peter using the same public word when they make their utterances. By requiring that lexical items be mapped onto one another in a one to one manner, the policy theorist would rule out that Peter and agent B are using the same word when they utter the sounds [pædərɛvski].

This position has some intuition behind it. When B articulates his lexical

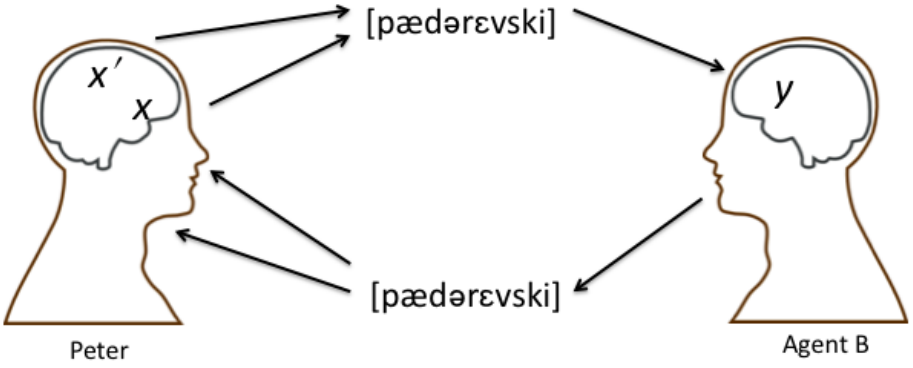

Figure 5.1: 'Paderewski' and Peter's Communicative Policies

item y, which of Peter's lexical items is the correct construal? Either? Neither? Both? I don't think the answer is straightforward. There may be no right answer as to which one of Peter's lexical items is the correct construal of B's utterance [pædərevski].

Furthermore, we might argue that it matters for communication that when Peter and B are construing each other's utterances, they are preserving representational sameness or distinctness. If Peter's and B's policies are as described by Figure 5.1, then B is construing two of Peter's utterances of distinct lexical items as the same, when they are not meant to be construed as the same. And Peter may be construing B's utterances of a single lexical item as different, when they are not meant to be construed as different. We might argue that this marks a

failure in understanding and a defect in their communicative exchange.

Imagine a community of linguistic agents who all have lexicons like Peter's. That is, they all mistakenly believe that there are two Paderewskis, one a talented musician and one a politician. Suppose this community exists a thousand years after Paderewski dies. Suppose for any two agents in this community, their lexical items for the Paderewski they think is a musician are mapped by their policies onto and only onto one another, and their lexical items for the Paderewski they think is a politician are mapped by their policies onto and only onto one another. In such a case, what impediment is there to saying there are two public words in use? Wouldn't we be failing to describe something important about their language and about their communicative exchanges by saying there is only one word (despite the fact that each agent's lexicon represents there being two)? And furthermore, doesn't communication succeed?

The alternative claim that there is a single public word available to Peter and B has some intuition behind it. Indeed, there is a single individual named once by his parents. But suppose a policy theorist claimed (as historical-intentionalists do) that Peter and B are using the same public word. It would seem that Peter's and B's policies look a lot like the policies in cases of homonyms (in which there fails to be a one to one mapping between lexical items). The policy theorist may then have to say that articulations of homonyms are articulations of the same public word. But surely this claim is controversial. Perhaps there is way to keep Paderewski cases separate from homonym cases, (after all, I've only argued for a necessary condition for articulating the same word) and perhaps a fleshed out policy view could achieve this. But I've yet to uncover it.

Question (11):

Sometimes historical origin (and not merely the conventions at work between agents) seems to make a difference to what public word is being used. Recall the case of two scientists A and B who independently discover the same species of tree, and independently name this species 'belm'. They then pass this name on to various individuals, including agent C (who acquires his lexical item x from A) and D (who acquires his lexical item y from B). We considered that C and D might meet in conversation, and coordinate their lexical items x and y via communicative policies. Here, it seemed explanatorily important to say that C and D are articulating the same word.

But suppose we find out that really A and B discovered different species. Doesn't it seem natural in this case to say that C and D were not using the same word, despite the fact that their communicative policies coordinated their lexical items?

Reply:

The policy theorist might say the following. If C and D's policies do in fact coordinate their lexical items, then it might be that C and D are using a single word 'belm' which refers to the two species of trees discovered by A and B. We might say that those agents have established a new public word, whose meaning was ambiguous between the two species that were discovered, or we might say that this new word refers to both species. (This might be in fact what happened with the English word "jade". There are in fact two kinds of jade, and many users of the English word had not known this, perhaps using a single word to refer ambiguously to both kinds.)

Question (12):

You haven't really said how to individuate public words. *Historical-Intentional Models* offer a clear answer: public words are individuated by their introduction events. What makes it the case that one public word is different from another on *The Communicative Policy Model*?

Reply:

I'll suggest that we may use sets of policies that coordinate lexical items as a basis for word-individuation. Such an option avoids some of the problems historical origin accounts face. In a sense, I've already gestured towards this idea, since the question how to individuate words and the question of when two utterances are utterances of the same word are closely related, if not two sides of the same issue.

First, to get the feel for why appealing to coordination and sets of communicative policies might be promising, consider an analogy with conventions. Recall that conventions are, according to Lewis, one kind of solution to a coordination problem. It is natural to think that two agents are participating in a single convention when they are performing their part of the adopted solution to a coordination problem. And it is natural to move from this thought to the thought that adopted solutions of a certain kind, or adopted pairs of actions that form solutions of the relevant kind, individuate conventions. It is natural to think that there are as many conventions as there are adopted solutions to coordination problems.

Sets of communicative policies also form solutions to a coordination problem posed by communication. And I've already argued that agents articulate the same word only when they they carry our their part of the set of policies forming the solution. From here, it might be natural to say there are as many public words as

there are adopted solutions to these communicative coordination problems—that the distinctions between public words just are the distinctions between different adopted solutions to these communicative coordination problem.

Of course, more work needs to be done to support this hypothesis, but we can already see how an account of the individuation of words with a coordinative flavor fairs better than the historical origin account. Recall some of the considerations against using historical-origin to individuate words.

First, we considered the scientist case, in which two agents C and D had lexical items with the same phonological profile and referent, but distinct historical origins. I suggested that once C and D started communicating with each other, the origins of their lexical items was irrelevant to whether they were articulating the same word. What seemed to matter instead was that the agents began treating their articulations as belonging to the same word. This example was meant to suggest that the historical-origin facts carve out words differently than the agents who use them and that this is problematic given the communicative role of words. At the very least, it seemed to deny that the distinctions between words used in a communicative exchange plays an important role in enabling communication.

Using sets of policies to individuate words might avoid this. *The Communicative Policy Model* claims that if C and D end up coordinating their lexical items via their communicative policies, there can be at most one public word in play, which they both use. We might avoid the situation in which two agents have coordinated their lexical items, but yet articulations of those lexical items belong to different words.

Second, we considered the related concern that if word-individuation is determined by historical origin, then it is possible that agents are, in general, ignorant of the distinctions between words used in a conversation they are party to. I

claimed that this is a detriment to the philosophical view of communication and the role of public words, as it seems to deny agents the ability to track which words are being used.

Using sets of communicative policies that coordinate lexical items as a basis for individuating words might avoid this criticism as well. If sets of policies carve out the distinctions between public words used in a communicative exchange, then agents might not be, in general, ignorant of these distinctions. As noted previously, the ability to track the distinctions between words would be merely dependent on the proper functioning of the agent's communicative competence. That is, so long as agents are carrying out their policies and acting in accordance with them, they will be accurately tracking the distinctions between words in their communicative exchange.

Recall that a Lewisan account of communication claims that agents can reliably communicate to the extent that they share knowledge of linguistic conventions. In a similar spirit, we may claim that agents can reliably communicate to the extent that they 'cognize' or can properly carry our their communicative policies.

The distinctions between words and an agent's ability to track these distinctions matter on traditional theories of linguistic communication. It is in part because I use the word 'awake' and not the word 'asleep' that my utterance carries the meaning that it does, and it is in part because you can reliably recover that I uttered the word 'awake' that you successfully understands me.

But historical accounts of word-individuation seem to undercut the utility of public words in communication by either (i) individuating words differently than communicating agents do or (ii) making those distinctions opaque to agents' cognitive capacities. Alternatively, looking to agents' dispositions for mapping lexical items onto phonetic forms and vice versa avoids these criticisms and draws

distinctions in explanatorily important ways. In light of these considerations, I suggest that we start replacing talk of origin events and causal chains with talk of coordination events and communicative networks.

5.3 Conclusion

A large part of this dissertation has been devoted to criticizing other views of words and I've highlighted where such views struggle to provide a model of public words as they function in linguistic communication. But we've actually ended up with a model of words that takes a fair amount of inspiration from previous views and which incorporates elements from both *Historical-Intentional Models* and *Internal-Lexical Models*.

The driving force behind Kaplan's *Historical-Intentional Model* is a rejection of *Replication Models*. In particular, he rejects the idea that a word is an eternal object whose utterances all sound or look alike. In its place, Kaplan offers a model on which a word is a created object, partially stored in the mental dictionaries of its users, whose utterances enjoy ample pronunciation variation. These ideas also animate *The Communicative Policy Model*. The idea that words are produced or built by the agents who use them and the sanctity of pronunciation variation are indeed part and parcel of my positive model. The two views part ways with respect to (i) the mechanism responsible for linking the entries in different agents' mental dictionaries and (ii) the mechanism responsible for linking utterances to the public word. But, the views are closer to each other than either of them is to *Replication Models*.

The Communicative Policy Model also incorporates aspects of *Internal-Lexical Models*, making use of phonological grammars and I-Languages to account for some kinds of pronunciation variation. It invokes a kind of Chomskyan compe-

tence in introducing communicative policies, which is broader than but interacting with the linguist's notion of linguistic competence. The two views part ways with respect to (i) the importance of social pressures on patterns of pronunciation and (ii) the possibility of a notion of a publicly shared structure used for communication.

Furthermore, *The Communicative Policy Model* finds inspiration in the Lewisian notions of a public language and a public word as objects used by a community as a matter of convention for the purposes of communication. The notion of a word on *The Communicative Policy Model* is fundamentally grounded in Lewis' notion of a solution to a coordination problem, and it sees the policies that act as solutions as akin to conventions. But, communicative policies differ from Lewisian conventions in that they need not be regularities (they may be in use only for a single communicative exchange), they need not be previously established (they may be established "on the fly"), and agents need not have explicit knowledge of them.

It might seem like we've ended up in a strange place. We started this dissertation with a view on which a word was akin to a simple string of sounds, and we've ended up with a view on which words are akin to sets of coordinated linguistic representations linked by networks of communicative policies. But given the complicated pronunciation variation facts, the demands on words in their communicative role, and the failure of other views to meet Chomsky's challenge, I'd say it is no surprise that public words are not the objects we once thought they were.

References

[1] J. Armstrong. *Language Change in Context*. PhD thesis, Rutgers University, 2013.

[2] A. Bergmann, K. Hall, and S. Ross, editors. *Language Files*. Ohio State University Press, 2007.

[3] S. Bromberger. What Are Words? comments on Kaplan (1990), on Hawthorne and Lepore, and on the issue. *Journal of Philosophy*, 108(9):486–503, 2011.

[4] R. Buchanan. A Puzzle About Meaning and Communication. *Nous*, 44(2):340–371, 2010.

[5] R. Buchanan. Reference, Understanding, and Communication. *Australasian Journal of Philosophy*, 92(1):55–70, 2014.

[6] H. Cappelen. Intentions in Words. *Nous*, 33(1):92–102, 1999.

[7] L. S. Casasanto. Using Social Information in Language Processing. *Berkeley Linguistics Society and the Linguistc Society of America*, 34(1), 2008.

[8] N. Chomsky. *Aspects of the Theory of Syntax*. MIT Press, 1965.

[9] N. Chomsky. *Knowledge of Language: Its Naure, Origin, and Use*. Praeger Publishers, 1986.

[10] N. Chomsky. *Language and Thought*. Moyer Bell, 1993.

[11] N. Chomsky. *New Horizons in the Study of Language and Mind*. Cambridge University Press, 2000.

[12] E. V. Clark and H. Clark. When Nouns Surface as Verbs. *Language*, 55(4):767–811, 1979.

[13] H. Clark. Making Sense of Nonce Sense. In G. B. Flores d'Arcais and R. J. Jarvella, editors, *The process of language understanding*, pages 297–331. Wiley, 1983.

[14] H. Clark. *Using Language*. Cambridge University Press, 1996.

[15] S. Cumming. Creatures of Darkness. *Analytic Philosophy*, 54(4):370–400, 2013.

[16] D. Davidson. Quotation. In *Inquires into Truth and Interpretation*. Oxford University Press, 1979.

[17] D. Davidson. A Nice Derangement of Epitaphs. In *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Blackwell Publishers, 1986.

[18] P. Foulkes, G. Docherty, and D. Watt. Phonological Variation in Child-Directed Speech. *Language*, 81(1):177–206, 2005.

[19] G. Frege. On Sense and Meaning. In Brian McGuinness, editor, *Frege: Collected Works*. Basil Blackwell, 1892.

[20] M. Gilbert. *On Social Facts*. Princton University Press, 1989.

[21] B. Gillon. Ambiguity, Inteterminancy, Deixis, and Vagueness. In *Semantics*. Oxford University Press, 2004.

[22] S. Haack. *Philosophy of Logic*. Cambridge University Press, 1978.

[23] J. Hawthorne and E. Lepore. On Words. *Journal of Philosophy*, 108(9):447–485, 2011.

[24] B. Hayes. *Introductory Phonology*. Blackwell Publishers, 2008.

[25] R. Kager. *Optimality Theory*. Cambridge University Press, 1999.

[26] D. Kaplan. Quantifying In. *Synthese*, 19(1/2):178–214, 1968.

[27] D. Kaplan. Words. *Proceedings of the Aristotelian Society*, 64:93–120, 1990.

[28] D. Kaplan. Words on Words. *The Journal of Philosophy*, 108(9):504–529, 2011.

[29] S. Kripke. A Puzzle About Belief. In A. Margalit, editor, *Meaning and Use*. Reidel, 1979.

[30] S. Kripke. *Naming and Necessity*. Harvard University Press, 1980.

[31] W. Labov. *Sociolinguistic Patterns*. University of Pennsylvania Press, 1972.

[32] E. Lepore and M. Stone. *Imagination and Convention*. Oxford University Press, 2015.

[33] D. Lewis. *Convention*. Harvard University Press, 1969.

[34] R. Millikan. *Language, Thought, and Other Biological Categories*. MIT Press, 1984.

[35] S. Pinker and R. Jackendoff. The faculty of language: what's special about it? *Cognition*, 95(2):201–236, 2005.

[36] W.V.O. Quine. *Mathematical Logic*. Cambridge University Press, 1951.

[37] W.V.O. Quine. Quantifiers and Propositional Attitudes. *Journal of Philosophy*, 53:177–187, 1956.

[38] M. Richard. *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge University Press, 1990.