

1 Introduction

Beyond words, spoken language involves prosody: intonation, loudness, timing, and the like. In conversation prosody is vital: it enables people to mark things as important or incidental, show interest in something or change the topic, be sympathetic or business-like, and so on. Without prosody, conversations would be just alternating short speeches: the human element would be lost.

This book explains how speakers of American English use prosody to accomplish things in conversation. While native speakers do this without conscious awareness, that does not mean it is simple. Attempts to pin down the details have faced many challenges, but now, in a remarkable convergence, researchers in diverse traditions – experimental phonetics, formal phonology, conversation analysis, and signal processing – have independently begun using compatible styles of description. The shared core is the notion of prosodic construction. Prosodic constructions are recurring temporal patterns of prosodic features that express specific meanings and functions. These typically involve not only intonation but also energy, speaking rate, timing, and articulation properties, often with synchronized contributions by two participants.

For example, consider one that is common in active listening. A listener can show interest and engagement by periodically nodding or saying *uh-huh* or the like, but this is not done at random. Example 1 illustrates.

1) A: *well, it's someone's house.*

B: *yeah*

A: *they're gonna, I mean, there's like, they're gonna be spinning.*

The audio for this and other examples is available at www.cambridge.org/ward.

In this example, B has asked what kind of party A is talking about. A is not entirely sure, so she starts telling what she knows. Importantly, her pitch, shown with the wavy line in Figure 1.1, goes down on *someone's* and stays low for a moment, then B backchannels with *yeah*.

2 1 Introduction



Figure 1.1 Pitch contour

	Speaker A	Speaker B
-600 ms	loudness increases	
-400 ms	pitch drops and stays low	
-300 ms	loudness decreases	
-100 ms	silence	
0 ms		backchannel, quiet, flat pitch
800 ms	resumes speaking, loud, fast, high pitch	

Figure 1.2 The Backchanneling Construction: A first approximation

This exemplifies a common pattern: a region of low pitch – typically below the speaker’s 26th percentile for at least 110 milliseconds – is followed by a short response by the listener about a half second later. This is usually part of a larger construction, as follows: in the middle of an utterance the speaker goes louder and articulates more clearly a word or two, here *someone’s*, then gently drops her pitch while reducing the energy, and then falls silent. The listener produces a backchannel, generally a quiet word in flat pitch, here *yeah*. The original speaker then says a few words with more energy, before resuming her normal speaking style. Figure 1.2 shows how these prosodic features are configured in time.

This pattern of interaction is not a meaningless ritual. It’s used when a speaker has a multi-part package of information to convey. Winding up one part, A gives B the chance to display attention, interest, and readiness to process the next part. After B does this with a backchannel, A continues. Only then, in many cases, does the speaker deliver the truly important information. In the example, A reveals that *they’re gonna be spinning*, implying that there will be dancing, only after the participants have jointly performed this prosodic pattern.

In casual conversation people typically use this pattern several times a minute, without awareness. In general, we use prosody without thinking about it. While we may struggle with the special rules for professional speaking – newsreading, political speaking, acting, and so on – like stressing the important words and pausing dramatically to underline important points, in fact, it is the prosody of conversation that is more complex. Rather than a few simple rules, there is a large set of resources to master. Effective speakers know these

forms and meanings and deftly select among them and combine them on the fly to exactly fit the situation and their goals. This seems easy only because these skills are acquired in childhood; as adults, we are never aware of all this, except when learning a new language.

While not the first book on the prosody of English, this book goes beyond previous work in several ways.

First, this book focuses on the uses of prosody in conversation. Many previous books have focused on monologue or on simple exchanges, describing the proper ways to say things like *Excuse me, do you have the time?* and *Would you like tea or coffee?* However, the prosody of such sentences is trivial compared to what happens when people go beyond such basic routines to actually explain things, make plans, cooperate on a task, or get to know each other. Among other things, speakers in dialog generally pursue multiple goals simultaneously, and prosody is important in this.

Second, this book covers all of prosody, not just intonation. While pitch is especially salient, many other aspects of prosody – loudness variations, syllable durations, other timing properties, and so on – are also essential for conveying many meanings.

Third, this book reports some new discoveries, including a key element of politeness in offers and invitations (late pitch peaks), a way we convey positive regard (a simple pitch-intensity configuration), and an essential characteristic of urgent warnings (breathy voice).

Fourth, this book strives to catalog all the most important prosodic patterns in American English conversation. This would have been impossible even a few years ago, but the notion of prosodic construction enables previously scattered findings to be connected into a comprehensive model. As we will see, much of the richness of prosody in conversation boils down to the uses of such constructions, plus one simple combining process: superposition.

Fifth, this book focuses on the facts, not issues of theory. While prosodic phenomena are relevant to many central questions in linguistics, the book will discuss these only in passing: the focus is on accurately describing the facts of English prosody in conversation.

Sixth, this book relates dialog behaviors to the larger picture of human interaction. People are inherently social creatures, and for thousands of years philosophers and scientists have striven to work out what this implies about human nature. Yet some of the most interesting evidence has been hiding in plain sight, in the nuts and bolts of how people interact, moment-by-moment, in everyday conversations. Thus this book will occasionally digress to speculate about connections to cultural values and practices: not modern attributes like popular foods and entertainment, but the fundamental, deeply rooted ways in which Americans manage social obligations, reach decisions, and connect with other people.

Seventh, this book is intended for anyone with an interest in language. No previous knowledge of prosody or phonetics is assumed, and the International Phonetic Alphabet (/aɪ.pi:.eɪ/) will not be used. The phenomena are described with a minimum of terminology and illustrated with many examples. These are crucial; to get maximum benefit from this book, please listen to the audio examples and then say each one out loud. If people are around, at least mouth them or speak them in your head. Prosody is processed using specific brain circuits (mostly in the right hemisphere, in contrast to words, which are processed on the left) and these are activated by auditory input. Just reading the examples may be interesting, but to fully experience the prosodic patterns you need to say them and hear them.

The book is in four parts. Chapters 2 through 4 lay the groundwork, each introducing a prosodic pattern and using it to explain fundamental concepts. Chapters 5 through 10 give historical perspective and explain how superposition-based modeling enables efficient discovery, interleaved with presentation of two more patterns. Then Chapters 11 through 13 catalog eighteen specific constructions involved in turn-taking, topic management, and expressing of stance, and finally, Chapters 14 and 15 give a wider perspective. We'll travel quickly, but I'll occasionally step off the path for anecdotes about research missteps and breakthroughs; for stories about the interesting prosody of children and foreigners; for speculations about culture, language learning, personality, social roles, emotion, politeness and rapport; and to debunk some myths.

Why should we care about the prosody of dialog? Most of the time we do just fine without thinking about it. But lack of awareness can cause problems. If we naively assume that prosody faithfully expresses intention and personality, someone who uses inappropriate prosody may seem impolite or insensitive, although the actual cause may be just a technical glitch in their pitch or timing, or a different idea about some pattern of English. If we instead see prosody as a system of its own, with its own logic and its own failure modes, we can reduce such misunderstandings.

I wrote this book out of frustration and hope. The frustration was from seeing valuable findings about conversational prosody go to waste, known and accessible only to scholars in a few isolated niches. At the same time non-native speakers of English, including many of my friends and co-workers, fail to communicate effectively because they never learned these constructions ... because they were never taught ... because their teachers lacked a resource on the facts of the prosody of dialog. My hope is that a clear account of prosodic constructions will help us all — scientists, teachers, co-workers, and friends — understand more deeply and communicate better.