Nonnative Prosody and the Intelligibility of Ambiguous Utterances

Nonnative Prosody and the Intelligibility of Ambiguous Utterances

Robin S. Schoener

A Thesis in the Field of Linguistics

For the Degree of Master of Liberal Arts in Extension Studies

Harvard University

November 2015

Abstract

This study examines nonnative prosody and intelligibility. Past research has suggested that prosody that is unfamiliar or inappropriate in some way can adversely affect the intelligibility of speech (e.g., Hahn, 2004; Tajima, Port & Dalby, 1997; Grover, Jamieson & Dobrovlosky, 1987; Field, 2005). In this study, the effect of overall prosody rather than the effects of particular prosodic features is analyzed.

Fifteen native and 15 nonnative speakers were recorded reading identical sets of ambiguous sentences while viewing cartoon drawings. Cartoons viewed by 8 members of each speaker group portrayed one of the two possible interpretations ("Version A") for each sentence. Cartoons seen by the remaining 7 speakers of each group showed the alternative ("Version B") interpretations. Recordings were divided and rearranged into new soundtracks containing a different speaker for every sentence. Fifteen native listeners viewed documents showing the Version A and Version B cartoons of each sentence side by side while listening to the new soundtracks, indicating which of the two cartoon versions they believed each speaker had viewed when recording.

Listeners identified the cartoon seen by the speaker significantly less often when the speaker was a nonnative, suggesting a relationship between speaker type and intelligibility. Results were further subdivided into 4 categories of structural ambiguity. Of those, compound noun vs. adjective + noun ambiguities (e.g. ***White*** *House* vs. *white* ***house***) accounted for most of listeners' errors in disambiguation.

Acknowledgments

Behind this endeavor there has been a group of mentors, colleagues, family and friends without whose support it would not have reached fruition, and I would like to express my sincerest gratitude to the many people who guided me, encouraged me, and offered help along the way.

I would first like to thank Dr. Kevin Ryan for agreeing to serve as my thesis director. His patient guidance and generousness with his time have been much appreciated. I am grateful as well to my research advisor, Dr. Dante Spetter, for her unceasing encouragement and unlimited patience.

My sincere thanks also to Dr. Susan Klimczak at the South End Technology Center as well as to Wendy Schoener at the University of Massachusetts, Boston, and to Susan Berry at Boston University's CELOP for their generous assistance in securing testing environments and locating potential participants. And as there could not have been a study without participants, I am indebted to all of them as well for their time and trust.

Finally, I would like to thank Hilda Spann, editor, supporter, but above all, friend.

Table of Contents

List of Tables

List of Figures

Chapter I

Introduction

Does the prosody of nonnative speech affect the intelligibility of ambiguous utterances?

Prosody, which is comprised of the intonation, rhythm, duration, and stress elements of speech, is used by speakers to communicate semantic, syntactic and pragmatic information (Carlson, 2009). Until relatively recently, the study of articulation, or segmentals, has outdistanced the study of prosody, or suprasegmentals, in speech intelligibility research. This imbalance has been particularly pronounced in research involving nonnative and disordered speech. The effect of prosody on intelligibility in pathological speech has been "marginalized and misunderstood" (Klopfenstein, 2009), while its role in nonnative speech has been "rarely examined" (Boula de Mareüil & Vieru-Dimulescu, 2006).

A growing body of research indicates that prosody in nonnative speech exerts a stronger negative effect on intelligibility than do other elements, such as syllable structure or segmentals (James, 1976; Anderson-Hsieh & Koehler, 1988; Anderson-Hsieh, Johnson & Koehler, 1992). Speech resynthesis has been used in recent studies to superimpose one or more features of nonnative prosody onto native speech. With some exceptions (Winters and O'Brien, 2013), nonnative or incongruous prosodic features have degraded intelligibility when mapped onto native speech (Field, 2005; Braun, Dainora & Ernestus, 2011; Tajima, Port and Dalby, 1997).

Munro & Derwing (1995a, 1995b) define intelligibility as "the extent to which a speaker's message is actually understood", a definition adopted by other researchers in the field (e.g. Hahn, 2004; Field, 2005; Rajadurai 2007; Kennedy & Trofimovitch 2008, Floccia, Butler, Goslin & Ellis 2009).

However, the tests used by Munro, Derwing and others have measured the intelligibility of only the segmental portion of nonnative speech. Tests have included measures such as counting the number of errors in listeners' responses to word- or phrase-matching tasks, or in tasks in which the listener transcribes phrases or sentences dictated by the nonnative speaker. While appropriate for measuring the extent to which a speaker's *words* have been understood, these measures reveal nothing about the speaker's success in transmitting his or her meaning.

In syntactically ambiguous sentences such as the following, segmentals alone cannot convey the speaker's precise meaning:

a) *Janis saw the boy with the binoculars*. (In this sentence, did *Janis* have the binoculars, or did *the boy* have the binoculars?)

b) *Susan and Terry or Joe will help.* (Are the alternatives in this sentence *Susan and Terry* vs. *Joe*, or are they *Susan and Terry* vs. *Susan and Joe*?)

c) *Sam and Susan's neighbor appeared on the deck.* (Did *Sam and a neighbor of Susan's* appear on the deck, or did *a neighbor of both Sam and Susan* appear on the deck?)

Clearly, only the prosodic element of speech has the potential to disambiguate the meaning of structures such as these. In this study, it is hypothesized that the prosody of nonnative speech affects the intelligibility of structurally ambiguous utterances. This

will be tested by comparing the number of successful outcomes in listener disambiguation of utterances produced by nonnative vs. native speakers.

Definition of Terms

Intelligibility: the extent to which a speaker's *words and meaning* have been correctly understood by a listener. In this study, intelligibility was measured by a binary score of correct vs. incorrect interpretation of the speaker's intended meaning.

Syntactic or structural ambiguity: multiple possible interpretations of a phrase or sentence due to syntactic structure. For example, *the French history teacher* is structurally ambiguous, since it is unclear whether the person referred to is *a teacher of French history* or *a French teacher of history*. Schafer, Speer, Warren & White (2000) investigated another type of structural ambiguity in examples of early vs. late closure structure, such as *When Hector calls Anna [comes running]* vs. *When Hector calls Anna [he begins to stutter]*.

Compound noun ambiguity: type of ambiguity in which a two-word phrase could be interpreted as either a compound noun or as a simple adjective + noun phrase. For example, though the words *hot plate* are often used to denote an electric appliance used for cooking and heating food (i.e., the compound noun ***hot*** *plate),* in other contexts, the two words could conceivably be expressed as an adjective + noun phrase to indicate an overly warm piece of dinnerware – a *hot* ***plate***.

Prepositional phrase attachment ambiguity: type of ambiguity as in the sentence The *man hit the boy with the umbrella*, in which it is unclear whether the prepositional phrase *with the umbrella* is connected with ("attached" to) *the boy* or *poked*.

And/or constituency ambiguity: ambiguity as in the phrase *She will call Sarah and Peter or Tom,* in which either *Sarah and Peter* or *Peter or Tom* form a larger constituent.

Broad/narrow possessive ambiguity: ambiguity of the type in the phrase *This is Joan and Susan's cousin,* in which it is unclear whether the *cousin* referred to is related to Joan and Susan both, or to Susan only.

Intonation contour: The contour of the fundamental frequency of the voice over an entire utterance.

Speaker: when capitalized where not normally required by conventional punctuation, refers specifically to the participants serving in the Speaker portion of this study (e.g. *nonnative Speakers, native Speakers*).

Listener: when capitalized where not normally required by conventional punctuation, refers specifically to the native English speakers who served as participants in the Listener portion of the study.

Background

Prosody is exploited by speakers to encode semantic, syntactic and pragmatic as well as emotional information (Carlson, 2009). Intonation alone can indicate whether a question is a suggestion, an objection, or a request for an action or information (Sag & Liberman, 1975). Sentence stress, the location of the greatest stress on a word or words in a sentence, is employed, among other things, to highlight the focus of the sentence, as in *I went to* ***Cuba*** *last year*. It can also highlight new vs. given information (***John*** *went to*

Cuba last year, too), and contrasts, as in *I ordered a **medium** coffee, not a **large*** (Akker & Cutler, 2003; Hahn, 2004).

Although there is a relationship between prosody and meaning, there is not a straightforward one-to-one correspondence between the two (Watson & Gibson, 2005; Carlson, Frazier & Clifton, Jr., 2009).

Some syntactic structures, such as the pause after parenthetical asides, for example (Carlson, Frazier & Clifton, Jr., 2009), require a particular prosodic expression. Conversely, while acknowledging that pragmatic intent does not predict prosody, Sag and Liberman (1975) identified a prosodic structure that was restricted to a single pragmatic usage. In other instances, prosodic structures are incompatible with particular syntactic structures or pragmatic usages (Sag & Liberman, 1975; Carlson, Frazier & Clifton, Jr., 2009).

In most cases, however, no specific prosodic structure is required and the speaker may choose considerably more freely (Frazier, Carlson & Clifton Jr., 2006). For a given syntactic structure, there is often not only an array of possible prosodic structures from which to choose, but a wide variability within as well as across speakers (Sag & Liberman, 1975; Schafer, Speer, Warren & White, 2000). In their study of intonational disambiguation, Schafer, Speer, Warren & White (2000) documented 25 distinct intonation patterns in 35 speech samples for a single syntactic structure, and 22 distinct patterns in 48 samples for another. Thus, prosody "tends to be influential rather than decisive in resolving syntactic ambiguity" (Carlson, 2009).

Speakers are not always conscious of their use of prosody (Watson & Gibson, 2005; Carlson, 2009). Some studies suggest that speakers employ disambiguating prosody even when the context is fully disambiguating. For example, it is unclear

whether the fragment *When Hector calls Anna…* is an instance of early or late closure until subsequent disambiguating content indicates that *calls* is used intransitively and *Anna* is the subject of a clause to follow (*When Hector calls, Anna comes running* – early closure) or whether *Anna* is the object of the transitive *calls* (*When Hector calls Anna, he begins to stutter* – late closure). Examining this type of ambiguity, Schafer, Speer, Warren & White (2000) recorded sentences spoken by participants during a cooperative game in which pairs of players with independent information attempted to move pieces around a board seen by only one of them. While freely uttered, the sentences of the player viewing the pieces were restricted to among three possibilities beginning with identical initial fragments, two of which were early closures and the other a late closure in the context of the entire utterance. When these identical initial fragments were then truncated from the recorded utterances, listeners were able to distinguish the early versus late closures on the basis of the prosodic information alone.

For their part, listeners monitor prosodic information (Schafer, Speer, Warren & White, 2000; Watson & Gibson, 2005), anticipating that speakers' prosody will signal the intended interpretation of their utterances (Carlson, Frazier, & Clifton, Jr., 2009).

Listeners exploit prosody not only to disambiguate syntax (Beach, 1991; Warren, Nolan, Grabe & Holst, 1995; Schafer, Speer, Warren & White, 2000; Carlson, Frazier & Clifton, Jr., 2009), but to predict upcoming sentence structure in real-time speech processing as well (Beach, 1991; Akker & Cutler, 2003). Like speakers, listeners are not always aware of their use of prosody (Beach, 1991), drawing on it to disambiguate syntax even when the ambiguity is only temporary (Fayer & Krasinski, 1987; Schafer, Speer, Warren & White, 2000; Carlson, Frazier, & Clifton, Jr., 2009).

Listeners also make use of prosody to resolve semantic ambiguities and discriminate among pragmatic interpretations. In certain types of cases this is achieved solely through intonation (Sag & Liberman, 1975; Mulac & Nash, 1977). Modifying pitch frequencies on a speech synthesizer to vary intonation while keeping all other aspects of prosody constant, Mulac and Nash (1977) showed results indicating that listeners were able to discriminate the semantic difference between *I thought so* meaning (*...and I was right)* vs. *I thought so...(but I was wrong*) by the speaker's intonation. Sag and Liberman (1975) also determined that listeners make use of intonation to disambiguate the pragmatic intention of speech acts, distinguishing, for example, a request or suggestion framed as a question, such as *Will you go to your room?* or *Why don't you sit next to George?* from a query for information (e.g., *Will you visit Rouen while you're in France?* or *Why don't you like pickles?*).

Misleading or incorrect prosodic information relayed by speakers can interfere with comprehension (Watson & Gibson, 2005). The effect of unfamiliar intonation alone can hamper listener processing. An unattested intonation contour developed in Dutch was found to slow lexical access in both resynthesized and naturally produced speech (Braun, Dainora & Ernestus (2011). Research has also shown that in short utterances, for example, violation of the "new vs. given information" sentence stress rule makes processing more difficult for the listener (Terken & Nooteboom, 1987; Hahn, 2004).

The development of prosodic sensitivity begins immediately after birth (Fernald & Kuhl, 1987). Infants can identify the prosody of their mother language within just a few months; for example, three-month-old American infants were shown to discriminate

among the prosody of English vs. Dutch or Italian (Schafer, Shucard, & Jaeger, 1999). Infants as young as four months demonstrate preferential attentiveness to motherese, the type of speech mothers use with their babies that is characterized by a markedly wider intonation range than that used in normal adult speech (Fernald & Kuhl, 1987).

Prosody, effectively used, can determine the intelligibility of ambiguous semantic content, syntactic structure and pragmatic intent, yet native-like prosody in a second language one of the last things nonnative speakers acquire (Pennington & Ellis, 2000). If they acquire it at all, "... most [nonnative] speakers end up with an intonational foreign accent... even if they are otherwise highly proficient second-language speakers..." (Braun, Dainora & Ernestus, 2011).

Studies of the effect of nonnative prosody on intelligibility have produced mixed results. In early research on the use of visual feedback for improving nonnative speech, James (1976) found evidence suggesting that nonnative prosody exerts a greater effect on intelligibility than articulation. Acceptability ratings were compared for nonnative Canadian French speakers of English whose prosody and articulation had earlier been separately rated. James found that a higher proportion of speakers with good prosody and poor articulation were judged as having better intelligibility than those with poor prosody but good articulation. However, as this was not the main focus of his study, statistical analysis was not performed.

Comparing the effects of inaccuracies in segmentals, prosody and syllable structure on expert raters' evaluations of nonnative intelligibility, Anderson-Hsieh, Johnson & Koehler (1992) found that while each of these variables influenced pronunciation ratings significantly, prosody had the strongest effect. Though the small

number of participants and the heterogeneity of the languages precluded strongly conclusive claims, outcomes were analyzed for indications of the relative effects of prosody, segmentals and syllable structure on intelligibility within the two main language subgroups. Results indicated that while the relative effects of these variables differed between the language subgroups, prosody exerted the strongest effect on intelligibility within both groups.

It is impossible to dissociate prosody in the natural speech stream from the effects of other aspects of speech that may influence intelligibility. In addition to articulation that is inaccurate or difficult to understand, factors such as listener familiarity with the specific nonnative accent in question, semantic content, speaking rate, and hesitations (Derwing & Munro, 1997, Gass & Varonis, 1984, Fayer & Krasinski, 1987, Anderson-Hsieh & Koehler, 1988) all may affect listeners' perceptions.

A stronger nonnative accent does not always result in less intelligible speech, however. The outcome of a study of the relationship between different aspects of Mandarin nonnative speech conducted by Munro & Derwing (1995b) did not show a significant relationship between accent and intelligibility. Listeners rated strength of accent of 30-second recorded slices in which nonnative Mandarin speakers freely described an amusing story illustrated in pictures. Even for speech rated as heavily accented, listeners were able to transcribe the recorded slices highly accurately.

The use of speech synthesizers has allowed researchers to dissociate prosody from other aspects of pronunciation by digitally manipulating features of the speech signal individually.

Using a synthesizer, Grover, Jamieson & Dobrovolsky (1987) manipulated French, German and English native speech samples, examining the effect of superimposing intonation onto native English and French utterances. Their results indicated that native listeners could not tell the difference between the intact native speech, and the resynthesized, "nonnative" speech. It is important to note, however, that they superimposed the intonation for only a single word. Several of the native English listeners also indicated they felt unable to distinguish the effect of intonation from the effect of stress.

Tajima, Port and Dalby (1997) modified and resynthesized utterances in English spoken by a native Chinese speaker and a native English speaker so that the durations of the native and nonnative speech were effectively switched – i.e., the segmental durations within short phrases in English spoken by the native Chinese speaker were resynthesized to match that of the same phrases spoken by the native English speaker, and vice versa. Participants took part in a forced-choice test in which they identified the recorded phrases from among three distractor phrases and the correct phrase. They found that the intelligibility of the Chinese nonnative speech was significantly improved, from 39% to 58%, by switching the duration to the native English duration pattern. That of native the English speech was significantly degraded, from 94% to 83%, when the durations were changed to the nonnative patterns. It should be noted that the phrases were two and three words in length.

Digitizing and resynthesizing the speech of native speakers, Winters and O'Brien (2013) transposed the duration, and in a separate test, the duration and intonation contour of German and English. A "SELF" resynthesis of the native prosody back onto the native

language for each group was created as a control for the effects of the resynthesis. The resynthesized utterances were then presented to native German and English speakers skilled in both languages, as well as to monolingual English speakers, in a sentence transcription task.

Contrary to what intuition might lead one to believe, native prosody mapped onto nonnative speech did not improve intelligibility, but instead rendered it less intelligible than the original nonnative utterance. Nonnative prosody mapped onto native speech was also less intelligible, although not as strikingly as in the Tajima, Port and Dalby (1997) study. Intelligibility for German among native German speakers dropped from 74.8% for the "SELF" control to 65.2% when the duration patterns and intonation were overlaid with English equivalents. Among native English speakers, intelligibility of English dropped from 74.1% for the "SELF" stimuli to 65% for stimuli with German duration patterns and intonation. In both cases, intelligibility was better when the listener shared the native language of the speaker. English and German have a different rhythmicity from Chinese, with English and German traditionally described as "stress-timed" and Chinese as "syllable-timed". The difference between these two types of timing may account in some measure for the less dramatic effect of prosody on intelligibility in this study than in the Tajima, Port and Dalby (1997) study.

Wholly synthesized speech, as well as that resynthesized by manipulating individual elements in natural speech, presents several issues to consider. Speech synthesizers do not always faithfully reproduce the speech in question (Tajima, Port & Dalby, 1997). Schafer, Speer, Warren & White (2000) acknowledged that by truncating the utterances under study, small but relevant parts at the ends of the fragments were lost.

Most importantly, because prosodic elements may interact with each other in natural speech (Field, 2005), manipulating a single prosodic element may create inaccurate stimuli (Winters & O'Brien, 2013).

Perhaps one of the biggest challenges in measuring intelligibility today is that the construct itself has been the object of imprecise and often conflicting definitions. In a variation of their definition cited earlier, Derwing and Munro (2005) defined *intelligibility* as "the extent to which a listener actually understands an utterance" and *comprehensibility* as "the listener's *perception* (emphasis mine) of how difficult it is to understand an utterance". In practice, however, these two terms have been used interchangeably by other researchers. Furthermore, *intelligibility* as defined and measured by Munro and Derwing has been restricted to intelligibility of articulation. Since *comprehensibility* under Munro and Derwing's definition, meanwhile, is a measure of the listener's perception, it is not a means by which to objectively measure the effect of the speech upon the listener.

The design of this study will help avoid some of the problems associated with intelligibility research. By allowing listeners to see the ambiguous text, then theoretically, prosody and articulation can be dissociated and the effect of nonnative articulation on intelligibility circumvented. Secondly, this design tests the prosody as it occurs in the natural speech stream, rather than testing a resynthesized feature or features prosody. Third, the actual effect on naive listeners, rather than their opinions of its effect, can be measured and compared.

Chapter II

Research Method

An experiment was conducted to comparing the intelligibility of nonnative speakers to native speakers in instances of ambiguous utterances.

Participants

Fifteen nonnative speakers of English were recruited to serve as participants in this study from among students in an intermediate-level class at Boston University's Center for English Language and Orientation Program (CELOP). Over half of these were from Saudi Arabia; the rest were from various countries (Appendix A).

Thirty native English speakers were recruited as participants from among personnel at the South End Technology Center at Tent City, Clarendon Residences in the Back Bay, the Copley branch of the Boston Public Library, and at Boston's Public Garden. Of these, 15 served as Speakers and 15 served as Listeners in the study. Former ESL teachers, linguistics students or teachers, trained musicians, and speakers of other languages were excluded.

Each participant was given a $5 gift card for Starbucks.

Materials

Two 15-item Speaker test series were created, "Version A" and "Version B". Each test item in Version A consisted of a cartoon accompanied by an accurate, but

structurally ambiguous caption, for example, *THE MAN POKED THE BOY WITH THE UMBRELLA*. The corresponding item in Version B consisted of the identical caption with a cartoon that illustrated the alternative interpretation. Captions were written in uppercase and without capitalization or punctuation in order to remove unwanted textual clues to aid in disambiguating meaning. Three practice cartoons preceded the test items. Captions for the practice cartoons were not ambiguous.

Documents were created and printed in which the Version A and Version B cartoons were illustrated side by side for each sentence. Each of the two carried the same capitalized, unpunctuated caption that the Speaker saw. Check boxes were included for the Listeners to indicate Version A, Version B or "Not Sure".

Procedure

Test items were shown to the Speakers on a MacBook Pro computer. Eight nonnative Speakers and 8 native Speakers were shown the Version A cartoons; the remaining 7 participants in each Speaker group were shown the Version B cartoons.

All Speakers were instructed to look at the cartoon, ensure they understood it, and then read the caption aloud. They were told that they could ask freely about objects in the cartoon they were unable to see or interpret, and nonnative Speakers were further instructed that before reading the text, they could ask for the meaning and pronunciation of any particular words in advance. Speakers' attention was not directed to the ambiguous nature of the text in the test materials for this study. Studies in which speakers were alerted to structural ambiguities beforehand resulted in increased listener accuracy in

identifying their meaning (Albritton, McKoon & Ratcliff, 1996) suggesting that the speakers' use of prosody was more deliberate than in natural, spontaneous speech.

Speakers were recorded on the internal microphone of the MacBook Pro. The researcher sat beside each Speaker, advancing the screen so that results were not affected by participant unfamiliarity with operating the computer.

Using Apple's Garage Band application, the recordings from each Speaker were divided into one-sentence segments corresponding each of the 15 test sentences. Segments were edited as follows: unprompted self-corrections of an entire sentence were accepted in place of the original. This occurred in 2 cases. Superfluous syllables at the beginning of the recorded sentences due to false starts, self-corrections, or in the case of one Speaker, a rather long *uhh* preceding each utterance, were removed in order to reduce distraction; there were several of these, mostly among the nonnative Speakers. Self-corrections made during or at the ends of test sentences were left intact, since editing these would require removing syllables from within the utterance.

At this point, 2 additional native English speakers were recruited and their recordings substituted for one missing segment, one severely degraded segment, and 9 segments in which a native Speaker inadvertently substituted words in the caption or revealed the interpretation of the cartoon by chuckling. There were no such problems with the nonnative Speakers' recordings.

The 450 individual sentence segments were then randomly rearranged and reassembled into 30 sound tracks of 15 segments each. To the extent needed, items were additionally manually rearranged so that not more than 3 native or nonnative Speakers would be heard consecutively. Each track contained one sample of each of the test

sentences. Fifteen 2-track pairs were created; each of these pairs contained one test item uttered by each of the 30 Speakers and was divided into a roughly equal number of Version A and Version B items. The entire 450-segment test set was arranged such that no Speaker was represented more than once per particular sentence (Appendix B). Segments were separated by 5 seconds of silence. Tracks were normalized to reduce the range of volume variation among Speakers.

Each Listener was given two copies of the Listener document, one for each track, and asked to listen to a pair of 15-item tracks. They were informed that the caption was a legitimate interpretation for either cartoon, and instructed to listen to each item and indicate which version of the corresponding cartoon they believed the Speaker was referring to, or "Not Sure", if appropriate. A critical assumption was made that because they could see the text, Listeners would ignore the effect of inaccurate or unusual pronunciation of phonemes, and focus only on the prosodic content of the spoken matter. Items were repeated at Listener request. The researcher operated playback of the soundtracks so that results would not be affected by the participants having to perform this added task. Listeners were allowed extra time to answer and proceed to the following page in their documents if the 5-second delay between segments was insufficient.

Participants listened to the recordings on a MacBook Pro computer with a Logitech 390 headset. A single practice item for Listeners consisting of two cartoons representing alternate interpretations of an ambiguous sentence preceded the test items. See Appendix C for Speaker and Listener practice and test items.

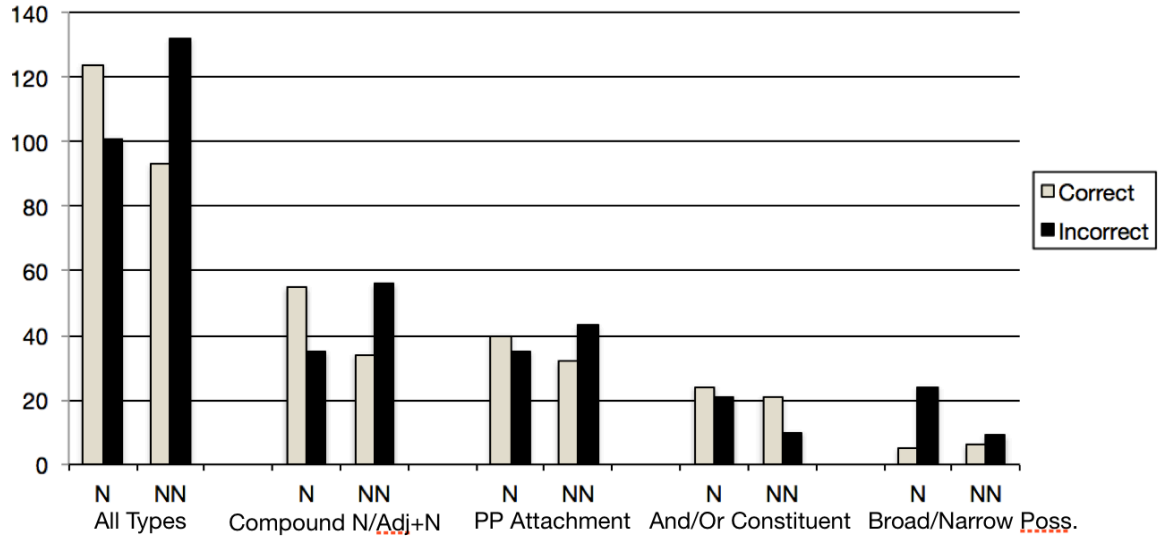
Chapter III

Results

Raw data for all test sentences combined showed that Listeners correctly identified the cartoons indicated by the native Speakers more often than those indicated by the nonnative Speakers (Figure 1). Aggregated by structural ambiguity type, data indicated that Listeners correctly identified the cartoons seen by Speakers more often when the Speaker was native than nonnative for compound noun and prepositional attachment ambiguities. In the case of and/or constituency, Listeners more often correctly identified the cartoon a Speaker had indicated whether the Speaker was native or nonnative; however, for the broad vs. narrow interpretation of the possessive in *Robert and Susan's niece*, they more often misidentified the appropriate cartoon, regardless of Speaker category.

Chi-square tests of independence were performed on each group of data to test the null hypothesis that there is no relationship between the Speaker category and Listeners' correct disambiguation of the Speakers' utterances. For all ambiguity types combined, the relationship between these variables was highly significant, $\chi^2 (1, 450) = 8.56$, $p < .01$ (Table 1). Odds ratios showed native Listeners were 0.57 times as likely to correctly interpret a structurally ambiguous utterance if the Speaker was nonnative. The effect size, at 0.138, was small.

Figure 1. Disambiguation by Ambiguity Type



Listener disambiguation of sentences by Speaker type and ambiguity type.

Table 1

Disambiguation – All Ambiguity Types

	Correctly Interpreted	Incorrectly Interpreted	Total
Native Speaker	124	101	225
Nonnative Speaker	93	132	225
Total	217	233	450

Listener disambiguation of sentences by Speaker type for all ambiguity types.

Of the individual types of ambiguity, Chi-square results indicated that the ability of native Listeners to correctly interpret compound noun ambiguities was significantly

related to whether the Speaker was native or nonnative, χ^2 (1, 180) = 9.80, p <.01. (Table 2). Odds ratios calculations showed that Listeners were 0.39 times as likely to correctly interpret this type of ambiguous utterance if the Speaker was nonnative. The p-value for this subgroup (p=.00344) was almost twice as high as for all results combined (p=.00174), and effect size, though still small, was larger than that of all results combined at 0.233, suggesting that the results for this category account for most of the overall effect.

Table 2

Disambiguation – Compound Noun Ambiguities.

	Correctly Interpreted	Incorrectly Interpreted	Total
Native Speaker	55	35	90
Nonnative Speaker	34	56	90
Total	89	91	180

Listener disambiguation of sentences by Speaker type for compound noun ambiguities.

Results of Chi-square tests for the other subgroups of data did not show a significant relationship between category of Speaker and Listener interpretation by type of ambiguity at either the p<.01 or p<.05 levels (Tables 3-5).

Table 3

Disambiguation – Prepositional Phrase Attachment Ambiguities

	Correctly Interpreted	Incorrectly Interpreted	Total
Native Speaker	40	35	75
Nonnative Speaker	32	43	75
Total	72	78	150

χ^2 (1, 150) = 1.7, p>.0.5 (not significant)

Listener disambiguation of sentences by Speaker type for prepositional phrase attachment ambiguities.

Table 4

Disambiguation – And/Or Constituency Ambiguities

	Correctly Interpreted	Incorrectly Interpreted	Total
Native Speaker	24	21	45
Nonnative Speaker	21	24	45
Total	45	45	90

χ^2 (1, 90) = 0.53, p>.0.5 (not significant)

Listener disambiguation of sentences by Speaker type for and/or constituency ambiguities.

Table 5

Disambiguation – Broad/Narrow Possessive Ambiguities

	Correctly Interpreted	Incorrectly Interpreted	Total
Native Speaker	5	10	15
Nonnative Speaker	6	9	15
Total	11	19	30

χ^2 (1, 30) = 0.71, p>.0.5 (not significant)

Listener disambiguation of sentences by Speaker type for broad/narrow possessive ambiguities.

Chapter IV

Discussion

The results of this study suggest that listeners may successfully disambiguate ambiguous syntactic structures more often if the speaker's prosody is native rather than nonnative. However, upon further analysis of the four types of ambiguity examined in this study, it was found that this outcome was determined almost exclusively by the compound noun ambiguity category. While the reasons for this are not definitively known, a few possibilities are considered here.

Listeners may have learned to detect smaller variations in some prosodic features than in others, in a process similar to the one by which we learn to perceive small phonemic contrasts as in own language, but as older children or adults often find ourselves at a loss to detect similar-sized contrasts in other languages. Therefore, the changes of pitch and stress required to distinguish a compound noun from its adjective + noun phrase isomorph may simply be more perceptible due to training than equivalent-sized changes in duration or rhythm. Alternatively, the size of the contrast produced by speakers in differentiating a left-stressed compound noun, as used here, from its adjective + noun phrase double may be objectively, measurably larger than that of features employed in disambiguating other structures, such as the change in duration that occurs at the ends of syntactic boundaries.

On the other hand, results in this category may be different because the two-word compound nouns used in this study can be, in effect, culturally determined vocabulary

items as well as structural items. While nonnative speakers of English might know, for example, that in English, the noun component of the adjective + noun phrase *red* ***light*** is generally stressed, they must learn that the compound noun ***red*** *light* is a particular type of red-colored light, i.e., the glowing red lantern in a traffic signal, or even the signal apparatus itself, regardless of which light is illuminated. They must know or learn that, in Boston, the ***Red*** *Line* is a particular subway route, indicated in red on Boston subway maps. They must also be aware that in certain well-known compound nouns such as ***hot*** *dog,* in which the compound is less obviously relatable to its individual parts, those individual constituents, under different circumstances, can be treated as the separate words of a simple adjective + noun phrase.

In an interesting development, it did not occur to a single native English Speaker viewing Version B of Sentence 2 that *BLACK BELT* was intended to be a compound noun, even though the girl pictured was clothed in karate garb; all of them pronounced the two words as adjective + noun, *black* ***belt.*** This underscores the effect of context on prosody. We generally hear the compound noun ***black*** *belt* in connection with the verbs *award, receive,* or *possess*: "*She has a black belt"*. It is less common to hear it connected to the concept of wearing, since it is not only a particular type of belt but it is connected to particular types of situations. If the caption had been *SHE HAS A BLACK BELT*, perhaps the more frequently used context would have had an effect on the results.

Generalizability

Though the outcome of this study suggests a relationship between prosody and intelligibility, the strength of that relationship is questionable and appears to be limited to

compound noun ambiguity. Generalizability is limited, because a convenience rather than a random sample of participants was used, and the nonnative participants consisted of students at a single level of English proficiency. The makeup of the nonnative sample was not controlled for languages or language subgroups. Results may vary widely for other nonnative languages such as tone languages, or may depend on the degree to which the nonnative language structure differs from English structures. Speakers may have been initially taught by a nonnative speaker with unusual prosody, or by a native speaker of a different dialect of English. Generalizability may also be affected to some extent by the particular dialect spoken by the native speaker and listener participants, as prosody can differ widely by national and regional dialect.

Comments on Design Issues

In studies of prosody in nonnative intelligibility, there will always be a tradeoff between utilizing non-synthesized speech and fully controlling for the effect of articulation; between the natural prosody of spontaneous speech and that of speech that is rehearsed or read from text; and between obtaining samples of extemporaneous speech and using matching samples of predetermined output that can be used in a speaker-to-speaker comparison.

The design of this study required the use of text, the prosody of which is not entirely the same as that of extemporaneous speech (Schafer, Speer, Warren & White, 2000). Therefore, the measure of how nonnative prosody affects intelligibility is limited to the context of the assigned text. In reading text, it must also be kept in mind that speakers are not simply the conveyers of the structure of their own utterances, but first

must interpret the structure of the text before them. Thus, they act in the roles of both receiver and producer of communication.

The intention of this study was to investigate the effect of prosody as an integrated phenomenon, separate from the effect of articulation, and to do so using matching sets of nonnative-native speaker utterances. Segmentals were "controlled" for by the use of the text. Utterance pairs could be matched by Speaker because they had text in front of them; Listeners knew what words Speakers were saying. Yet structural properties in the texts allowed for two possible interpretations, which the Listeners could only disambiguate by prosody, if at all.

Because participants were aware they were taking part in a language study, they tended to be rather cautious; nonnative Speakers seemed particularly focused on pronunciation, while some of the native Speakers found the captions so simple that they suspected and tried to uncover an imagined ruse.

Based on a casual side view of Speakers' eye movements as the cartoons were advanced on the screen, it appeared that at times both native, and especially nonnative, Speakers were focusing immediately on the text, reading it before looking at the corresponding cartoon. On this suspicion, recruitment of an entirely new group of nonnative Speaker participants was undertaken at the University of Massachusetts. This new group of nonnative Speakers was shown the cartoon first without the caption, then again with the caption added. There appeared to be no difference in the way the second nonnative group performed than the first, and this effort was abandoned and recording with the original Speakers was resumed.

Even when Speakers appeared to be viewing the cartoon before reading the caption, a default interpretation of the text often seemed to override the obvious interpretation of the visual cues, particularly in the case of compound noun ambiguity. For example, Version A of Item 5 clearly showed three canines. Yet even several of the native Speakers read the words *HOT DOGS* as ***hot*** *dogs,* the compound noun**,** some audibly chuckling at the presumed joke as they did so. Viewing Version B of Item 10, one of the nonnative Speakers asked in an aside, "Is that the **White** House? Oh, I want to visit the **White** House!" with appropriate pitch and stress both times. As he proceeded to read the caption, however, he read the words *WHITE HOUSE* as the adjective + noun phrase, *white* ***house***.

The cartoons themselves were occasionally the source of confusion, although this was often not presented to the researcher until after the recordings were finished. For example, at least one Speaker failed to understand that the purpose of the circles inside the "think bubbles" was to indicate which two objects inside the bubble should be treated a unit.

Problems with the size of the items to be viewed arose during the recordings of native English-speaking participants for the speaker series, as their ages covered a much wider range than did the ages of the nonnative Speaker participants. A few items had to be enlarged to such an extent for one native Speaker that, in effect, the cartoon was viewed separately from the caption.

Conclusion and Suggestions for Future Study

More study of natural speech prosody is needed. Although the use of synthesizers is becoming increasingly popular in comparing the effect of prosody on speech intelligibility, their use has been limited to shorter utterances and one or two prosodic features at a time. The entire prosody of one language cannot simply be overlaid onto another using a synthesizer. Furthermore, as Field (2005) noted, prosody interacts with segmentals.

Prosody in extemporaneous speech is ideal, but it is difficult to capture data that is specific and equivalent to use in a study comparing the effect of native vs. nonnative prosody on listeners. The study by Schafer, Speer, Warren & White (2000) offers a starting point. Short, memorizable sentences with ambiguous structures could be created and tested in an interactive game. As an alternative, an improvement on the design in this study might be made by having speakers describe the cartoon without reciting captions, after a discussion, or using a questions-and-answer technique. Data could be used directly if comparable, or truncated if necessary.

There appears to be a need for more study with respect to preferential analyses of structure due to functions in natural speech processing. If the Minimal Attachment Theory (Frazier, 1979) provides an accurate framework for describing the way that we naturally process real-time syntactic ambiguity, for example, then a bias in native speakers' responses to such ambiguity is implied that must be accounted for in studies of intelligibility. An interesting approach would be to compare the number of times participants correctly disambiguate structure in terms of the alternatives: do they have a tendency to prefer one structural analysis over another? Do they correctly analyze the

structure of a sentence in which the prepositional phrase is attached to the verb more often than when it is attached to the object of the verb, regardless of prosody? Does this phenomenon influence native and nonnative participants differently?

It is not impossible to find examples of situations in which, of a group of native listeners addressed by a speaker of their own dialect, one or more fail to correctly understand the words or meaning of the intended message. Yet, aside from *comprehensibility,* which is a perceptual measure, no method has been established that quantifies the listener's contribution to intelligibility. It would be interesting for future studies to examine intelligibility by the relative contribution made in terms of the speaker and listener. Clearly, successful communication is not entirely the responsibility of the speaker. It is the result of productive interaction between speaker and listener.

Appendix A

Speaker Participant Demographic Data

Nonnative English Speakers by Region, Age

	Country	Gender	Age
1	Saudi Arabia	M	18
2	Saudi Arabia	F	19
3	Saudi Arabia	M	19
4	Saudi Arabia	M	19
5	Saudi Arabia	M	19
6	Saudi Arabia	M	19
7	Saudi Arabia	M	25
8	Saudi Arabia	M	34
9	Qatar	M	22
10	Kazakhstan	M	37
11	Republic of Korea	F	22
12	People's Republic of China	F	22
13	People's Republic of China	F	26
14	Spain	F	19
15	Brazil	M	24

Median Age: 22

Native English Speakers by Age

	Gender	Age
1	M	21
2	F	21
3	F	26
4	M	29
5	M	36
6	M	46
7	F	56
8	F	57
9	M	64
10	F	65
11	F	67
12	F	67
13	F	70
14	F	82
15	M	86

Median Age: 57

Appendix B

Test Design

Listener-Track	Recorded Segment 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
L1 -1	N 1.1	NN 13.2	N 2.3	NN 1.4	N 8.5	N 4.6	NN 15.7	N 5.8	NN 11.9	N 11.10	NN 12.11	N 7.12	N 15.13	NN 6.14	N 12.15
L1 -2	NN 3.1	N 6.2	NN 2.3	N 3.4	NN 10.5	N 9.6	N 10.7	NN 5.8	NN 9.9	NN 4.10	N 14.11	NN 8.12	NN 14.13	N 13.14	NN 7.15
L2 -1	NN 6.1	N 12.2	N 1.3	NN 13.4	N 2.5	NN 1.6	N 8.7	N 4.8	NN 15.9	N 5.10	NN 11.11	N 11.12	NN 12.13	N 7.14	N 15.15
L2-1	N 13.1	NN 7.2	NN 3.3	N 6.4	NN 2.5	N 3.6	NN 10.7	N 9.8	N 10.9	NN 5.10	NN 9.11	NN 4.12	N 14.13	NN 8.14	NN 14.15
L3-1	N 15.1	NN 6.2	N 12.3	N 1.4	NN 13.5	N 2.6	NN 1.7	N 8.8	N 4.9	NN 15.10	N 5.11	NN 11.12	N 11.13	NN 12.14	N 7.15
L3-2	NN 14.1	N 13.2	NN 7.3	NN 3.4	N 6.5	NN 2.6	N 3.7	NN 10.8	N 9.9	N 10.10	NN 5.11	NN 9.12	NN 4.13	N 14.14	NN 8.15
L4-1	N 7.1	N 15.2	NN 6.3	N 12.4	N 1.5	NN 13.6	N 2.7	NN 1.8	N 8.9	N 4.10	NN 15.11	N 5.12	NN 11.13	N 11.14	NN 12.15
L4-2	NN 8.1	NN 14.2	N 13.3	NN 7.4	NN 3.5	N 6.6	NN 2.7	N 3.8	NN 10.9	N 9.10	N 10.11	NN 5.12	NN 9.13	NN 4.14	N 14.15
L5-1	NN 12.1	N 7.2	N 15.3	NN 6.4	N 12.5	N 1.6	NN 13.7	N 2.8	NN 1.9	N 8.10	N 4.11	NN 15.12	N 5.13	NN 11.14	N 11.15
L5-2	N 14.1	NN 8.2	NN 14.3	N 13.4	NN 7.5	NN 3.6	N 6.7	NN 2.8	N 3.9	NN 10.10	N 9.11	N 10.12	NN 5.13	NN 9.14	NN 4.15
L6-1	N 11.1	NN 12.2	N 7.3	N 15.4	NN 6.5	N 12.6	N 1.7	NN 13.8	N 2.9	NN 1.10	N 8.11	N 4.12	NN 15.13	N 5.14	NN 11.15
L6-2	NN 4.1	N 14.2	NN 8.3	NN 14.4	N 13.5	NN 7.6	NN 3.7	N 6.8	NN 2.9	N 3.10	NN 10.11	N 9.12	N 10.13	NN 5.14	NN 9.15
L7-1	NN 11.1	N 11.2	NN 12.3	N 7.4	N 15.5	NN 6.6	N 12.7	N 1.8	NN 13.9	N 2.10	NN 1.11	N 8.12	N 4.13	NN 15.14	N 5.15
L7-2	NN 9.1	NN 4.2	N 14.3	NN 8.4	NN 14.5	N 13.6	NN 7.7	NN 3.8	N 6.9	NN 2.10	N 3.11	NN 10.12	N 9.13	N 10.14	NN 5.15
L8-1	N 5.1	NN 11.2	N 11.3	NN 12.4	N 7.5	N 15.6	NN 6.7	N 12.8	N 1.9	NN 13.10	N 2.11	NN 1.12	N 8.13	N 4.14	NN 15.15
L8-2	NN 5.1	NN 9.2	NN 4.3	N 14.4	NN 8.5	NN 14.6	N 13.7	NN 7.8	NN 3.9	N 6.10	NN 2.11	N 3.12	NN 10.13	N 9.14	N 10.15
L9-1	NN 15.1	N 5.2	NN 11.3	N 11.4	NN 12.5	N 7.6	N 15.7	NN 6.8	N 12.9	N 1.10	NN 13.11	N 2.12	NN 1.13	N 8.14	N 4.15
L9-2	N 10.1	NN 5.2	NN 9.3	NN 4.4	N 14.5	NN 8.6	NN 14.7	N 13.8	NN 7.9	NN 3.10	N 6.11	NN 2.12	N 3.13	NN 10.14	N 9.15
L10-1	N 4.1	NN 15.2	N 5.3	NN 11.4	N 11.5	NN 12.6	N 7.7	N 15.8	NN 6.9	N 12.10	N 1.11	NN 13.12	N 2.13	NN 1.14	N 8.15
L10-2	N 9.1	N 10.2	NN 5.3	NN 9.4	NN 4.5	N 14.6	NN 8.7	NN 14.8	N 13.9	NN 7.10	NN 3.11	N 6.12	NN 2.13	N 3.14	NN 10.15
L11-1	N 8.1	N 4.2	NN 15.3	N 5.4	NN 11.5	N 11.6	NN 12.7	N 7.8	N 15.9	NN 6.10	N 12.11	N 1.12	NN 13.13	N 2.14	NN 1.15
L11-2	NN 10.1	N 9.2	N 10.3	NN 5.4	NN 9.5	NN 4.6	N 14.7	NN 8.8	NN 14.9	N 13.10	NN 7.11	NN 3.12	N 6.13	NN 2.14	N 3.15
L12-1	NN 1.1	N 8.2	N 4.3	NN 15.4	N 5.5	NN 11.6	N 11.7	NN 12.8	N 7.9	N 15.10	NN 6.11	N 12.12	N 1.13	NN 13.14	N 2.15
L12-2	N 3.1	NN 10.2	N 9.3	N 10.4	NN 5.5	NN 9.6	NN 4.7	N 14.8	NN 8.9	NN 14.10	N 13.11	NN 7.12	NN 3.13	N 6.14	NN 2.15
L13-1	N 2.1	NN 1.2	N 8.3	N 4.4	NN 15.5	N 5.6	NN 11.7	N 11.8	NN 12.9	N 7.10	N 15.11	NN 6.12	N 12.13	N 1.14	NN 13.15
L13-1	NN 2.1	N 3.2	NN 10.3	N 9.4	N 10.5	NN 5.6	NN 9.7	NN 4.8	N 14.9	NN 8.10	NN 14.11	N 13.12	NN 7.13	NN 3.14	N 6.15
L14-1	NN 13.1	N 2.2	NN 1.3	N 8.4	N 4.5	NN 15.6	N 5.7	NN 11.8	N 11.9	NN 12.10	N 7.11	N 15.12	NN 6.13	N 12.14	N 1.15
L14-2	N 6.1	NN 2.2	N 3.3	NN 10.4	N 9.5	N 10.6	NN 5.7	NN 9.8	NN 4.9	N 14.10	NN 8.11	NN 14.12	N 13.13	NN 7.14	NN 3.15
L15-1	N 12.1	N 1.2	NN 13.3	N 2.4	NN 1.5	N 8.6	N 4.7	NN 15.8	N 5.9	NN 11.10	N 11.11	NN 12.12	N 7.13	N 15.14	NN 6.15
L15-2	NN 7.1	NN 3.2	N 6.3	NN 2.4	N 3.5	NN 10.6	N 9.7	**N** 10.8	NN 5.9	NN 9.10	NN 4.11	N 14.12	NN 8.13	NN 14.14	N 13.15

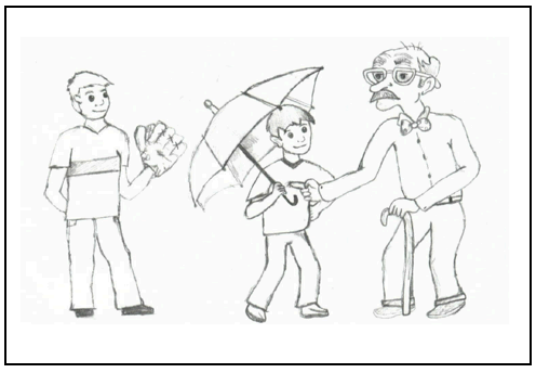
N - Native Speaker, NN – Nonnative Speaker Speaker number precedes decimal; Sentence number follows decimal

Appendix C

Practice and Test Sentences for Ambiguous Structures

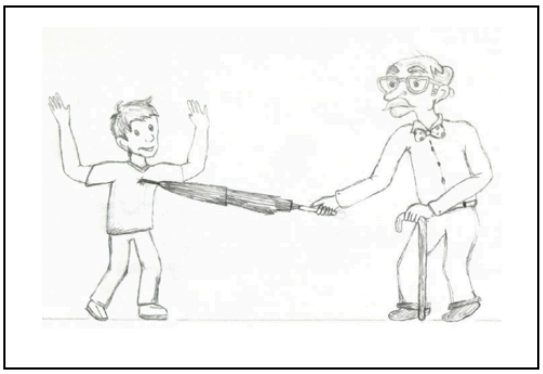
Some native and nonnative Speakers recorded the caption while viewing the Version A cartoon; the rest recorded while viewing Version B. Listeners heard the recordings while viewing both versions, and indicated which one they thought the Speaker had viewed.

Version A



THE MAN POKED THE BOY WITH THE UMBRELLA.

Version B



THE MAN POKED THE BOY WITH THE UMBRELLA.

Speaker Practice Sentences

Sentence 1. HE IS A TRUCK DRIVER.
Pictured: A man standing in front of a tractor-trailer cab.

Sentence 2. YOU MAY HAVE COFFEE OR TEA.
Pictured: A server carries a tray with 2 cups, one labeled *coffee*, the other *tea*.

Sentence 3. PUT THE BOOK ON THE TABLE.
Pictured: A book floats midair. Nearby is a table. An arrow points from the book to the tabletop.

Listener Practice Sentence. SHE IS A FRENCH TEACHER.

Version A: A teacher stands in front of a blackboard with "Basic French" and French-English vocabulary is written on it.

Version B: A teacher stands in front of a blackboard with "La Velocité" and some equations written on it. The Eiffel tower is seen through the window.

Test Sentences

Sentence 1. DAN SAW THE RED LIGHT.

Version A: A driver in a car sees a traffic light ahead of him. The light is red.

Version B: A driver in a car sees a tower topped by a red light in the distance.

Sentence 2. THE GIRL IS WEARING A BLACK BELT.

Version A: A girl is wearing street clothes accessorized with a black belt.

Version B: A girl is wearing karate uniform with black belt.

Sentence 3. THE MAN POKED THE BOY WITH THE UMBRELLA.

Version A: An old man pokes one of two boys in the chest – the boy holding an open umbrella.

Version B: An old man leaning on a cane pokes a boy in the chest using the tip of his furled umbrella,

Sentence 4. THIS IS ROBERT AND SUSAN'S NIECE.

Version A: A picture of a man and a woman.

Version B: A picture of a woman.

Sentence 5. THESE ARE THREE HOT DOGS.

Version A: Three canines are panting on a beach under a blazing sun.

Version B: Three frankfurters in buns are topped with condiments.

Sentence 6. I WILL BRING ICE CREAM AND COOKIES OR CAKE.

Version A: A boy is shown with a "speak bubble" overhead. Inside it are a tub of ice cream and, circled together, some cookies and a cake.

Version B: A boy is shown with a "speak bubble" overhead. Inside it are a tub of ice cream and some cookies, circled together, and, separately, a cake.

Sentence 7. THE LADIES' ROOM IS MESSY.

Version A: Women's clothes and shoes are strewn all over a rumpled bedroom.

Version B: A women's lavatory with overflowing wastebasket and toilet paper strewn on the floor.

Sentence 8. SHE POINTED TO THE RED LINE.

Version A: Of a red triangle, red circle and red line pictured on a blackboard, a woman points to the line.

Version B: A woman in front of the MBTA map of Boston points to the Red Line.

Sentence 9. THE MAN SPOKE TO THE WOMAN IN THE CAR.

Version A: A man and a woman stand beside a convertible. The man speaks to a woman sitting in the convertible.

Version B: A man and a woman sitting in a car are having a conversation.

Sentence 10. THERE IS A FENCE IN FRONT OF THE WHITE HOUSE.

Version A: A white house with a fence in front.

Version B: The White House with a fence in front.

Sentence 11. YOU MUST SHOW A CREDIT CARD AND A LICENSE OR A PASSPORT.

Version A: A man is shown with a "speak bubble" overhead. Inside it are a credit card and a license, circled together, and, separately, a passport.

Version B: A man is shown with a "speak bubble" overhead. Inside it are a credit card and, circled together, a license and a passport.

Sentence 12. PUT THE HAT ON THE DOG IN THE BOX.

Version A: Two dogs are pictured; one peeks out of a box. An arrow points to from a hat to the dog peeking out of the box.

Version B: A cat and a dog, both with hats, and an empty box are pictured. An arrow points from the dog's hat to the box.

Sentence 13. THE WOMAN REACHED OUT TO THE GIRL WITH A SMILE.

Version A: A woman reaches out to shake hands with a girl. The woman is smiling.

Version B: A woman reaches out to shake hands with one of two girls – the girl who is smiling.

Sentence 14. PUT THE MONEY ON THE TABLE NEXT TO THE LAMP.

Version A: Currency floats midair. Nearby, a table is pictured, with a book and a lamp on it. An arrow points from the currency to a spot next to the lamp.

Version B: Currency floats midair between two tables, one of which is next to a lamp. An arrow points from the currency to a spot on the table located next to the lamp.

Sentence 15. SHE WILL INVITE MARY AND STEVE OR TOM.

Version A: A girl is shown with a "think bubble" overhead. Inside it are a girl labeled MARY and a boy labeled STEVE, circled together, and, separately, a boy labeled TOM.

Version B: A girl is shown with a "think bubble" overhead. Inside it are a girl labeled MARY and, circled together, a boy labeled STEVE and another labeled TOM.

Appendix D

Raw Data

Data is ordered by Speaker Type, Ambiguity Type, and Sentence ID

Event	Sentence ID	Ambiguity Type	Speaker ID	Speaker Type	Image Seen	Listener ID	Listener Response	Result
16	1	comp.noun	3	N	A	1	b	0
76	1	comp.noun	14	N	B	3	b	1
106	1	comp.noun	8	N	B	4	b	1
121	1	comp.noun	12	N	B	5	a	0
166	1	comp.noun	4	N	B	6	n	0
181	1	comp.noun	11	N	A	7	a	1
196	1	comp.noun	9	N	A	7	a	1
226	1	comp.noun	5	N	A	8	a	1
241	1	comp.noun	15	N	A	9	a	1
316	1	comp.noun	10	N	B	11	n	0
331	1	comp.noun	1	N	A	12	a	1
376	1	comp.noun	2	N	B	13	a	0
391	1	comp.noun	13	N	A	14	a	1
436	1	comp.noun	7	N	A	15	a	1
406	1	comp.noun	16	N	B	14	a	0
17	2	comp.noun	16	N	B	1	a	0
32	2	comp.noun	12	N	B	2	a	0
77	2	comp.noun	13	N	A	3	a	1
92	2	comp.noun	15	N	A	4	a	1
122	2	comp.noun	7	N	A	5	a	1
167	2	comp.noun	14	N	B	6	n	0
182	2	comp.noun	11	N	A	7	a	1
242	2	comp.noun	5	N	A	9	a	1
287	2	comp.noun	10	N	B	10	a	0
302	2	comp.noun	4	N	B	11	a	0
317	2	comp.noun	9	N	A	11	a	1
332	2	comp.noun	8	N	B	12	b	1
377	2	comp.noun	3	N	A	13	b	0
392	2	comp.noun	2	N	B	14	a	0
422	2	comp.noun	1	N	A	15	b	0
5	5	comp.noun	8	N	B	1	b	1
35	5	comp.noun	2	N	B	2	b	1

80	5	comp.noun	16	N	B	3	b	1
95	5	comp.noun	1	N	A	4	a	1
125	5	comp.noun	12	N	B	5	b	1
170	5	comp.noun	13	N	A	6	a	1
185	5	comp.noun	15	N	A	7	b	0
215	5	comp.noun	7	N	A	8	a	1
260	5	comp.noun	14	N	B	9	b	1
275	5	comp.noun	11	N	A	10	a	1
335	5	comp.noun	5	N	A	12	b	0
380	5	comp.noun	10	N	B	13	b	1
395	5	comp.noun	4	N	B	14	b	1
410	5	comp.noun	9	N	A	14	a	1
440	5	comp.noun	3	N	A	15	b	0
22	7	comp.noun	10	N	B	1	a	0
37	7	comp.noun	8	N	B	2	b	1
82	7	comp.noun	3	N	A	3	a	1
97	7	comp.noun	2	N	B	4	a	0
142	7	comp.noun	16	N	B	5	a	0
157	7	comp.noun	1	N	A	6	a	1
187	7	comp.noun	12	N	B	7	a	0
232	7	comp.noun	13	N	A	8	a	1
247	7	comp.noun	15	N	A	9	a	1
277	7	comp.noun	7	N	A	10	b	0
322	7	comp.noun	14	N	B	11	n	0
337	7	comp.noun	11	N	A	12	a	1
397	7	comp.noun	5	N	A	14	a	1
427	7	comp.noun	4	N	B	15	b	1
442	7	comp.noun	9	N	A	15	a	1
8	8	comp.noun	5	N	A	1	a	1
38	8	comp.noun	4	N	B	2	b	1
53	8	comp.noun	9	N	A	2	a	1
68	8	comp.noun	8	N	B	3	a	0
113	8	comp.noun	3	N	A	4	a	1
128	8	comp.noun	2	N	B	5	b	1
173	8	comp.noun	16	N	B	6	a	0
188	8	comp.noun	1	N	A	7	b	0
218	8	comp.noun	12	N	B	8	b	1
263	8	comp.noun	13	N	A	9	b	0
278	8	comp.noun	15	N	A	10	b	0
308	8	comp.noun	7	N	A	11	a	1
353	8	comp.noun	14	N	B	12	a	0

368	8	comp.noun	11	N	A	13	a	1
443	8	comp.noun	10	N	B	15	b	1
10	10	comp.noun	11	N	A	1	a	1
40	10	comp.noun	5	N	A	2	a	1
85	10	comp.noun	10	N	B	3	a	0
100	10	comp.noun	4	N	B	4	a	0
115	10	comp.noun	9	N	A	4	b	0
130	10	comp.noun	8	N	B	5	b	1
175	10	comp.noun	3	N	A	6	b	0
190	10	comp.noun	2	N	B	7	b	1
235	10	comp.noun	16	N	B	8	b	1
250	10	comp.noun	1	N	A	9	a	1
280	10	comp.noun	12	N	B	10	b	1
325	10	comp.noun	13	N	A	11	b	0
340	10	comp.noun	15	N	A	12	b	0
370	10	comp.noun	7	N	A	13	a	1
415	10	comp.noun	14	N	B	14	b	1
3	3	prep phr.	2	N	B	1	b	1
33	3	prep phr.	1	N	A	2	b	0
63	3	prep phr.	12	N	B	3	b	1
108	3	prep phr.	13	N	A	4	a	1
123	3	prep phr.	15	N	A	5	b	0
153	3	prep phr.	7	N	A	6	a	1
198	3	prep phr.	14	N	B	7	b	1
213	3	prep phr.	11	N	A	8	a	1
273	3	prep phr.	5	N	A	10	b	0
318	3	prep phr.	10	N	B	11	b	1
333	3	prep phr.	4	N	B	12	b	1
348	3	prep phr.	9	N	A	12	b	0
363	3	prep phr.	8	N	B	13	b	1
408	3	prep phr.	3	N	A	14	b	0
438	3	prep phr.	16	N	B	15	b	1
54	9	prep phr.	10	N	B	2	a	0
69	9	prep phr.	4	N	B	3	a	0
84	9	prep phr.	9	N	A	3	b	0
99	9	prep phr.	8	N	B	4	b	1
144	9	prep phr.	3	N	A	5	b	0
159	9	prep phr.	2	N	B	6	b	1
204	9	prep phr.	16	N	B	7	b	1
219	9	prep phr.	1	N	A	8	a	1
249	9	prep phr.	12	N	B	9	a	0

294	9	prep phr.	13	N	A	10	a	1
309	9	prep phr.	15	N	A	11	a	1
339	9	prep phr.	7	N	A	12	a	1
384	9	prep phr.	14	N	B	13	a	0
399	9	prep phr.	11	N	A	14	a	1
429	9	prep phr.	5	N	A	15	b	0
12	12	prep phr.	7	N	A	1	a	1
42	12	prep phr.	11	N	A	2	a	1
102	12	prep phr.	5	N	A	4	a	1
147	12	prep phr.	10	N	B	5	a	0
162	12	prep phr.	4	N	B	6	a	0
177	12	prep phr.	9	N	A	6	a	1
192	12	prep phr.	8	N	B	7	a	0
237	12	prep phr.	3	N	A	8	a	1
252	12	prep phr.	2	N	B	9	a	0
297	12	prep phr.	16	N	B	10	a	0
312	12	prep phr.	1	N	A	11	a	1
342	12	prep phr.	12	N	B	12	a	0
387	12	prep phr.	13	N	A	13	a	1
402	12	prep phr.	15	N	A	14	a	1
447	12	prep phr.	14	N	B	15	a	0
13	13	prep phr.	15	N	A	1	a	1
58	13	prep phr.	14	N	B	2	a	0
73	13	prep phr.	11	N	A	3	a	1
133	13	prep phr.	5	N	A	5	a	1
178	13	prep phr.	10	N	B	6	n	0
193	13	prep phr.	4	N	B	7	a	0
208	13	prep phr.	9	N	A	7	a	1
223	13	prep phr.	8	N	B	8	a	0
268	13	prep phr.	3	N	A	9	a	1
283	13	prep phr.	2	N	B	10	b	1
328	13	prep phr.	16	N	B	11	b	1
343	13	prep phr.	1	N	A	12	b	0
373	13	prep phr.	12	N	B	13	a	0
418	13	prep phr.	13	N	A	14	a	1
433	13	prep phr.	7	N	A	15	b	0
29	14	prep phr.	13	N	A	1	a	1
44	14	prep phr.	7	N	A	2	b	0
89	14	prep phr.	14	N	B	3	a	0
104	14	prep phr.	11	N	A	4	b	0
164	14	prep phr.	5	N	A	6	b	0

209	14	prep phr.	10	N	B	7	a	0
224	14	prep phr.	4	N	B	8	b	1
239	14	prep phr.	9	N	A	8	a	1
254	14	prep phr.	8	N	B	9	b	1
299	14	prep phr.	3	N	A	10	a	1
314	14	prep phr.	2	N	B	11	a	0
359	14	prep phr.	16	N	B	12	a	0
374	14	prep phr.	1	N	A	13	n	0
404	14	prep phr.	12	N	B	14	a	0
434	14	prep phr.	15	N	A	15	a	1
6	6	and/or	4	N	B	1	a	0
21	6	and/or	9	N	A	1	b	0
51	6	and/or	3	N	A	2	b	0
66	6	and/or	2	N	B	3	b	1
111	6	and/or	16	N	B	4	a	0
126	6	and/or	1	N	A	5	n	0
156	6	and/or	12	N	B	6	b	1
201	6	and/or	13	N	A	7	a	1
216	6	and/or	15	N	A	8	b	0
246	6	and/or	7	N	A	9	b	0
291	6	and/or	14	N	B	10	b	1
306	6	and/or	11	N	A	11	b	0
366	6	and/or	5	N	A	13	b	0
411	6	and/or	10	N	B	14	b	1
426	6	and/or	8	N	B	15	b	1
26	11	and/or	14	N	B	1	a	0
71	11	and/or	5	N	A	3	a	1
116	11	and/or	10	N	B	4	a	0
131	11	and/or	4	N	B	5	b	1
146	11	and/or	9	N	A	5	a	1
161	11	and/or	8	N	B	6	b	1
206	11	and/or	3	N	A	7	a	1
221	11	and/or	2	N	B	8	b	1
266	11	and/or	16	N	B	9	b	1
281	11	and/or	1	N	A	10	a	1
311	11	and/or	12	N	B	11	a	0
356	11	and/or	13	N	A	12	b	0
371	11	and/or	15	N	A	13	b	0
401	11	and/or	7	N	A	14	a	1
431	11	and/or	11	N	A	15	a	1
15	15	and/or	12	N	B	1	a	0

45	15	and/or	15	N	A	2	b	0
75	15	and/or	7	N	A	3	b	0
120	15	and/or	14	N	B	4	a	0
135	15	and/or	11	N	A	5	a	1
195	15	and/or	5	N	A	7	a	1
240	15	and/or	10	N	B	8	a	0
255	15	and/or	4	N	B	9	b	1
270	15	and/or	9	N	A	9	a	1
285	15	and/or	8	N	B	10	b	1
330	15	and/or	3	N	A	11	a	1
345	15	and/or	2	N	B	12	a	0
390	15	and/or	16	N	B	13	a	0
405	15	and/or	1	N	A	14	a	1
450	15	and/or	13	N	A	15	a	1
19	4	posess.	3	N	A	1	b	0
49	4	posess.	16	N	B	2	a	0
64	4	posess.	1	N	A	3	b	0
94	4	posess.	12	N	B	4	b	1
139	4	posess.	13	N	A	5	a	1
154	4	posess.	15	N	A	6	a	1
184	4	posess.	7	N	A	7	b	0
229	4	posess.	14	N	B	8	b	1
244	4	posess.	11	N	A	9	b	0
304	4	posess.	5	N	A	11	b	0
349	4	posess.	10	N	B	12	a	0
364	4	posess.	4	N	B	13	n	0
379	4	posess.	9	N	A	13	b	0
394	4	posess.	8	N	B	14	b	1
424	4	posess.	2	N	B	15	a	0
31	1	comp.noun	6	NN	B	2	n	0
1	1	comp.noun	1	NN	A	1	b	0
46	1	comp.noun	13	NN	A	2	b	0
61	1	comp.noun	15	NN	A	3	a	1
91	1	comp.noun	7	NN	A	4	a	1
136	1	comp.noun	14	NN	B	5	b	1
151	1	comp.noun	11	NN	A	6	n	0
211	1	comp.noun	5	NN	A	8	b	0
256	1	comp.noun	10	NN	B	9	a	0
271	1	comp.noun	4	NN	B	10	b	1
286	1	comp.noun	9	NN	A	10	b	0
301	1	comp.noun	8	NN	B	11	b	1

346	1	comp.noun	3	NN	A	12	a	1
361	1	comp.noun	2	NN	B	13	b	1
421	1	comp.noun	12	NN	B	15	a	0
2	2	comp.noun	13	NN	A	1	a	1
47	2	comp.noun	7	NN	A	2	b	0
62	2	comp.noun	6	NN	B	3	a	0
107	2	comp.noun	14	NN	B	4	a	0
137	2	comp.noun	8	NN	B	5	a	0
152	2	comp.noun	12	NN	B	6	a	0
197	2	comp.noun	4	NN	B	7	a	0
212	2	comp.noun	11	NN	A	8	a	1
227	2	comp.noun	9	NN	A	8	a	1
257	2	comp.noun	5	NN	A	9	b	0
272	2	comp.noun	15	NN	A	10	a	1
347	2	comp.noun	10	NN	B	12	a	0
362	2	comp.noun	1	NN	A	13	a	1
407	2	comp.noun	2	NN	B	14	b	1
437	2	comp.noun	3	NN	A	15	b	0
20	5	comp.noun	10	NN	B	1	a	0
50	5	comp.noun	2	NN	B	2	a	0
65	5	comp.noun	13	NN	A	3	b	0
110	5	comp.noun	3	NN	A	4	b	0
140	5	comp.noun	7	NN	A	5	b	0
155	5	comp.noun	6	NN	B	6	b	1
200	5	comp.noun	14	NN	B	7	b	1
230	5	comp.noun	8	NN	B	8	b	1
245	5	comp.noun	12	NN	B	9	a	0
290	5	comp.noun	4	NN	B	10	b	1
305	5	comp.noun	11	NN	A	11	a	1
320	5	comp.noun	9	NN	A	11	b	0
350	5	comp.noun	5	NN	A	12	b	0
365	5	comp.noun	15	NN	A	13	n	0
425	5	comp.noun	1	NN	A	15	b	0
7	7	comp.noun	15	NN	A	1	b	0
52	7	comp.noun	10	NN	B	2	a	0
67	7	comp.noun	1	NN	A	3	b	0
112	7	comp.noun	2	NN	B	4	b	1
127	7	comp.noun	13	NN	A	5	b	0
172	7	comp.noun	3	NN	A	6	b	0
202	7	comp.noun	7	NN	A	7	a	1
217	7	comp.noun	6	NN	B	8	a	0

262	7	comp.noun	14	NN	B	9	a	0
292	7	comp.noun	8	NN	B	10	a	0
307	7	comp.noun	12	NN	B	11	a	0
352	7	comp.noun	4	NN	B	12	b	1
367	7	comp.noun	11	NN	A	13	a	1
382	7	comp.noun	9	NN	A	13	n	0
412	7	comp.noun	5	NN	A	14	a	1
23	8	comp.noun	5	NN	A	1	a	1
83	8	comp.noun	10	NN	B	3	a	0
98	8	comp.noun	1	NN	A	4	a	1
143	8	comp.noun	2	NN	B	5	a	0
158	8	comp.noun	13	NN	A	6	n	0
203	8	comp.noun	3	NN	A	7	a	1
233	8	comp.noun	7	NN	A	8	b	0
248	8	comp.noun	6	NN	B	9	a	0
293	8	comp.noun	14	NN	B	10	a	0
323	8	comp.noun	8	NN	B	11	b	1
338	8	comp.noun	12	NN	B	12	b	1
383	8	comp.noun	4	NN	B	13	b	1
398	8	comp.noun	11	NN	A	14	a	1
413	8	comp.noun	9	NN	A	14	a	1
428	8	comp.noun	15	NN	A	15	b	0
25	10	comp.noun	4	NN	B	1	b	1
55	10	comp.noun	5	NN	A	2	b	0
70	10	comp.noun	15	NN	A	3	b	0
145	10	comp.noun	10	NN	B	5	a	0
160	10	comp.noun	1	NN	A	6	b	0
205	10	comp.noun	2	NN	B	7	a	0
220	10	comp.noun	13	NN	A	8	b	0
265	10	comp.noun	3	NN	A	9	b	0
295	10	comp.noun	7	NN	A	10	b	0
310	10	comp.noun	6	NN	B	11	a	0
355	10	comp.noun	14	NN	B	12	a	0
385	10	comp.noun	8	NN	B	13	n	0
400	10	comp.noun	12	NN	B	14	a	0
430	10	comp.noun	11	NN	A	15	a	1
445	10	comp.noun	9	NN	A	15	a	1
18	3	prep phr.	2	NN	B	1	b	1
48	3	prep phr.	3	NN	A	2	b	0
78	3	prep phr.	7	NN	A	3	b	0
93	3	prep phr.	6	NN	B	4	b	1

138	3	prep phr.	14	NN	B	5	a	0
168	3	prep phr.	8	NN	B	6	n	0
183	3	prep phr.	12	NN	B	7	b	1
228	3	prep phr.	4	NN	B	8	b	1
243	3	prep phr.	11	NN	A	9	b	0
258	3	prep phr.	9	NN	A	9	b	0
288	3	prep phr.	5	NN	A	10	b	0
303	3	prep phr.	15	NN	A	11	b	0
378	3	prep phr.	10	NN	B	13	b	1
393	3	prep phr.	1	NN	A	14	b	0
423	3	prep phr.	13	NN	A	15	b	0
9	9	prep phr.	11	NN	A	1	a	1
24	9	prep phr.	9	NN	A	1	a	1
39	9	prep phr.	15	NN	A	2	a	1
114	9	prep phr.	10	NN	B	4	a	0
129	9	prep phr.	1	NN	A	5	a	1
174	9	prep phr.	2	NN	B	6	a	0
189	9	prep phr.	13	NN	A	7	a	1
234	9	prep phr.	3	NN	A	8	a	1
264	9	prep phr.	7	NN	A	9	b	0
279	9	prep phr.	6	NN	B	10	b	1
324	9	prep phr.	14	NN	B	11	a	0
354	9	prep phr.	8	NN	B	12	b	1
369	9	prep phr.	12	NN	B	13	n	0
414	9	prep phr.	4	NN	B	14	a	0
444	9	prep phr.	5	NN	A	15	b	0
27	12	prep phr.	8	NN	B	1	b	1
57	12	prep phr.	4	NN	B	2	n	0
72	12	prep phr.	11	NN	A	3	a	1
87	12	prep phr.	9	NN	A	3	a	1
117	12	prep phr.	5	NN	A	4	a	1
132	12	prep phr.	15	NN	A	5	n	0
207	12	prep phr.	10	NN	B	7	a	0
222	12	prep phr.	1	NN	A	8	a	1
267	12	prep phr.	2	NN	B	9	a	0
282	12	prep phr.	13	NN	A	10	a	1
327	12	prep phr.	3	NN	A	11	b	0
357	12	prep phr.	7	NN	A	12	a	1
372	12	prep phr.	6	NN	B	13	a	0
417	12	prep phr.	14	NN	B	14	a	0
432	12	prep phr.	12	NN	B	15	a	0

28	13	prep phr.	14	NN	B	1	b	1
43	13	prep phr.	12	NN	B	2	n	0
88	13	prep phr.	4	NN	B	3	a	0
103	13	prep phr.	11	NN	A	4	a	1
118	13	prep phr.	9	NN	A	4	n	0
148	13	prep phr.	5	NN	A	5	a	1
163	13	prep phr.	15	NN	A	6	n	0
238	13	prep phr.	10	NN	B	8	a	0
253	13	prep phr.	1	NN	A	9	b	0
298	13	prep phr.	2	NN	B	10	a	0
313	13	prep phr.	13	NN	A	11	a	1
358	13	prep phr.	3	NN	A	12	b	0
388	13	prep phr.	7	NN	A	13	b	0
403	13	prep phr.	6	NN	B	14	a	0
448	13	prep phr.	8	NN	B	15	b	1
14	14	prep phr.	6	NN	B	1	a	0
59	14	prep phr.	8	NN	B	2	a	0
74	14	prep phr.	12	NN	B	3	a	0
119	14	prep phr.	4	NN	B	4	a	0
134	14	prep phr.	11	NN	A	5	a	1
149	14	prep phr.	9	NN	A	5	a	1
179	14	prep phr.	5	NN	A	6	a	1
194	14	prep phr.	15	NN	A	7	a	1
269	14	prep phr.	10	NN	B	9	a	0
284	14	prep phr.	1	NN	A	10	a	1
329	14	prep phr.	2	NN	B	11	a	0
344	14	prep phr.	13	NN	A	12	b	0
389	14	prep phr.	3	NN	A	13	a	1
419	14	prep phr.	7	NN	A	14	a	1
449	14	prep phr.	14	NN	B	15	a	0
36	6	and/or	1	NN	A	2	a	1
81	6	and/or	2	NN	B	3	a	0
96	6	and/or	13	NN	A	4	b	0
141	6	and/or	3	NN	A	5	b	0
171	6	and/or	7	NN	A	6	b	0
186	6	and/or	6	NN	B	7	b	1
231	6	and/or	14	NN	B	8	b	1
261	6	and/or	8	NN	B	9	a	0
276	6	and/or	12	NN	B	10	b	1
321	6	and/or	4	NN	B	11	a	0
336	6	and/or	11	NN	A	12	b	0

351	6	and/or	9	NN	A	12	a	1
381	6	and/or	5	NN	A	13	b	0
396	6	and/or	15	NN	A	14	a	1
441	6	and/or	10	NN	B	15	b	1
11	11	and/or	12	NN	B	1	b	1
41	11	and/or	11	NN	A	2	a	1
56	11	and/or	9	NN	A	2	a	1
86	11	and/or	5	NN	A	3	b	0
101	11	and/or	15	NN	A	4	b	0
176	11	and/or	10	NN	B	6	a	0
191	11	and/or	1	NN	A	7	b	0
236	11	and/or	2	NN	B	8	b	1
251	11	and/or	13	NN	A	9	a	1
296	11	and/or	3	NN	A	10	b	0
326	11	and/or	7	NN	A	11	a	1
341	11	and/or	6	NN	B	12	b	1
386	11	and/or	14	NN	B	13	a	0
416	11	and/or	8	NN	B	14	a	0
446	11	and/or	4	NN	B	15	a	0
30	15	and/or	7	NN	A	1	a	1
60	15	and/or	14	NN	B	2	n	0
90	15	and/or	8	NN	B	3	b	1
105	15	and/or	12	NN	B	4	a	0
150	15	and/or	4	NN	B	5	n	0
165	15	and/or	11	NN	A	6	a	1
180	15	and/or	9	NN	A	6	a	1
210	15	and/or	5	NN	A	7	b	0
225	15	and/or	15	NN	A	8	a	1
300	15	and/or	10	NN	B	10	a	0
315	15	and/or	1	NN	A	11	b	0
360	15	and/or	2	NN	B	12	a	0
375	15	and/or	13	NN	A	13	a	1
420	15	and/or	3	NN	A	14	a	1
435	15	and/or	6	NN	B	15	a	0
4	4	posess.	1	NN	A	1	a	1
34	4	posess.	13	NN	A	2	b	0
79	4	posess.	3	NN	A	3	a	1
109	4	posess.	7	NN	A	4	a	1
124	4	posess.	6	NN	B	5	b	1
169	4	posess.	14	NN	B	6	n	0
199	4	posess.	8	NN	B	7	a	0

214	4	posess.	12	NN	B	8	a	0
259	4	posess.	4	NN	B	9	b	1
274	4	posess.	11	NN	A	10	b	0
289	4	posess.	9	NN	A	10	b	0
319	4	posess.	5	NN	A	11	a	1
334	4	posess.	15	NN	A	12	b	0
409	4	posess.	10	NN	B	14	a	0
439	4	posess.	2	NN	B	15	a	0

References

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition, 6*(2), 81-96.

Albritton, D., Mckoon, G., & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(3), 714-735.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*, 529-555.

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38*(4), 561-613.

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language, 30*(6), 644-663.

Boula de Mareüil, P., & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica, 63*, 247-267.

Braun, B., Dainora, A. & Ernestus, M. (2011). An unfamiliar intonation contour slows down online speech comprehension. *Language and Cognitive Processes, 26*(3), 350-375.

Carlson, K. (2009). How prosody influences sentence comprehension. *Language and Linguistics Compass, 3*(5), 1188-1200.

Carlson, K., Frazier, L., & Clifton, C., Jr. (2009). How prosody constrains comprehension: A limited effect of prosodic packaging. *Lingua, 119*(7), 1066-1082.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19*(1), 1-16.

Fayer, J. M., & Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning, 37*(3), 313-326.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*, 279-293.

Field, J. (2005). Intelligibility and the Listener: The Role of Lexical Stress. *TESOL Quarterly*, *39*(3), 399-423.

Frazier, L. (1979). *On comprehending sentences: syntactic parsing strategies* (Doctoral dissertation.) Available from ProQuest Dissertations and Theses database (UMI No. 7914150)

Frazier, L., Carlson, K., Clifton, Jr., C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, 10(6), 244-249.

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and Foreign Accent Processing in English: Can Listeners Adapt? *Journal of Psycholinguistics Research*, 38, 379-412.

Gass, S. M., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning, 34*, 65-89.

Grover, C., Jamieson, D. G., & Dobrovolsky, M. B. (1987). Intonation in English, French and German: Perception and production. *Language and Speech, 30*(3), 277-295.

Hahn, L. D., (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *Tesol Quarterly, 8*(2), 201-233.

James, E. F. (1976). The acquisition of prosodic features of speech using a speech visualizer. *IRAL, International Review of Applied Linguistics in Language Teaching, 14*(3), 227-243.

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review, 64*(3), 459-489.

Klopfenstein, M. (2009). Interaction between prosody and intelligibility. *International Journal of Speech-Language Pathology, 11*(4), 326-331.

Léon, P. R., Martin, P. (1972). Applied Linguistics and the Teaching of Intonation, *The Modern Language Journal*,56(3),139-144.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289-306.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49*(supplement 1), 285-310.

Derwing, T, & Munro, M. J. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly: A Journal for Teachers of English to Speakers of Other Languages and of Standard English as a Second*

Dialect, *39*(3), 379-397.

Mulac, A., & Nash, R. (1977). Effects of intonation pattern of synthesized and natural speech upon listener resolution of semantic ambiguity. *Language & Speech, 20*(3), 274-279.

Pennington, N. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal, 84*(3), 372-389.

Rajadurai, J. (2007). Intelligibility studies: A consideration of empirical and ideological issues. *World Englishes, 26*(1), 87-98.

Sag, I. A., & Liberman, M. (1975). The intonational disambiguation of indirect speech acts. *Papers from the Regional Meetings, Chicago Linguistic Society, 11*, 487-497.

Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research, 29*(2), 169-182.

Shafer, V.L., Shucard, D.W. & Jaeger, J.J. (1999). Electrophysiological indices of cerebral specialization and the role of prosody in language acquisition in 3-month-old infants. *Developmental Neuropsychology, 15*(1), 73-109.

Terken, J, & Nooteboom, S.G. (1987)**.** Opposite effects of accentuation and deaccentuation on verification latencies for given and new information, *Language and Cognitive Processes, 2*(3-4), 145-163.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, 25, 1-24.

Warren, P., Nolan, F., Grabe, E., & Holst, T (1995). Post-lexical and prosodic

phonological processing, *Language and Cognitive Processes, 10*(3-4), 411-417

Watson, D. & Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica* 59(2-3), 279-300.

Winters, S. & O'Brien, M. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication, 55*(201), 486-507.