Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter know, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Robert James Smith April 18th, 2011

Examining the role of prosody in the resolution of semantic ambiguity in L1 and L2 speakers of English

by

Robert James Smith

Marjorie Pak, Ph.D.
Adviser

Program in Linguistics

Marjorie Pak, Ph.D.
Adviser

Susan Tamasi, Ph. D.
Committee Member

Judy Raggi Moore, Ph.D.
Committee Member

April 18th, 2011

Examining the role of prosody in the resolution of semantic ambiguity
in L1 and L2 speakers of English

by

Robert James Smith

Marjorie Pak, Ph.D.
Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Linguistics

2011

Abstract
Examining the role of prosody in the resolution of semantic ambiguity
in L1 and L2 speakers of English

By Robert James Smith

Semantic ambiguity arises when utterances contain linguistic constructions that allow for multiple logical meanings. Speech element that yield such ambiguity include the focus-sensitive operators *only* and *even*, the quantifier *every*, and the negation *not* when paired with *because.* In linguistics, prosody refers to the musical, rhythmic quality of speech that allows speakers to package units of language into intelligible phrase structures. This project attempted to identify the prosodic pitch changes associated with resolving instances of semantic ambiguity in native and non-native speakers of English. Forty-eight participants were recruited, sixteen each of English, Korean, and Spanish L1s. Participants in the production study were recorded as they read through a series of 20 short paragraphs, each containing an unmarked, semantically ambiguous utterance. The recordings were then analyzed using the software *Praat*. In the perception portion of this study, participants were asked to listen to audio clips of semantically ambiguous utterances and identify the corresponding target paragraph. Production analysis suggests that native English speakers varied their pitch changes to disambiguate between the two interpretations of the focus operators, but not the "every" and "not-because" contexts. The non-native English speakers did not systematically resolve any of contexts through pitch changes. In the perception study, native English speakers correctly selected target interpretations of the utterances at an overall rate of 76% ($p<0.05$). Performance varied greatly by each of the four types of ambiguity. The L2 English speakers performed significantly worse than the native English speakers. The results reveal that English speakers use prosody as a supplementary mechanism for resolving semantic ambiguity; however, second language learners do not readily acquire or recognize these prosodic patterns.

Examining the role of prosody in the resolution of semantic ambiguity in L1 and L2 speakers of English

By

Robert James Smith

Marjorie Pak, Ph.D.
Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors
Program in Linguistics

2011

Acknowledgements

I would like to thank my thesis advisor, Dr. Marjorie Pak, for her earnest support, ceaseless patience, and encouraging laughter; the entire Program in Linguistics, for introducing me to the wonderful world of language; and Scholarly Inquiry & Research at Emory, for helping fund my research.

Table of Contents

List of Figures

List of Tables

I. Introduction

In linguistics, the study of semantics encompasses the study of meaning. The concept of semantic ambiguity arises when linguistic constructions yield multiple, often contrary, meanings. As a result, the English language essentially maintains a staple of unresolved ambiguity within its semantic system. The theoretical motivation for this project stems from my desire to understand how it is possible for a vibrant, complex entity such as language to contain this inherent element of ambiguity when one of its primary functions is communicative. I am enthralled by the speaker's capacity to successfully wade through the mires of language with such little hesitation. What linguistic mechanisms do we use to unravel the nuance of meaning? Are speakers cognizant of these mechanisms? Are these cues universal or language-specific? If the latter, are they perceived and acquired by non-native speakers? All of these questions aim to explore the nature of ambiguity.

Of course, the potential explanations for the phenomenon of multiple meanings are limitless. For the purpose of developing a structured approach to these questions, I have chosen to focus on one particular tool that speakers have at their disposal to produce and perceive language: prosody. Simply described, prosody is the musical quality of speech that is both tandem to, yet distinct from, language's phonetic elements. It is the rhythm and pitches that make language comprehensible (Jun 1996: xvi). Individual functional units of language ("words" or "morphemes") are composed of phonemes ("psychological sound units") that sequence together to yield meaning. The relationship between the specific combination of sounds and the meaning that the combination represents is arbitrary. In other words, there is no inherent meaning attached to individual speech sounds, such as [d], [a], or [g]. Speakers then pattern these morphemes together within the framework of a defined syntax to produce intelligible phrases. The direct,

logical, denotative meaning that results from this sequencing of sounds, words, and phrases refers to semantics. In combination, these various levels comprise the core structural levels of language. However, none of these various nuts and bolts of language address the way, manner, and style with which language is expressed. Those elements form the basis for understanding the concept of prosody.

III.A. Prosody

Key terms that often fall under the scope of the term "prosody" include pitch, accent, stress, duration, rhythm, and tone, among many others. The role of prosody as a mechanism for grouping informational packets of language has long been a topic of study (Jun 1996: 3). These intonational facets of human speech can provide significant information regarding the typology of an utterance (i.e. whether a statement is interrogative or declarative), the illocutionary force of an utterance (i.e. the speaker's implied or intended meaning), the emotional state of the speaker, the emphasis or focus of an utterance, the existence of new vs. given information, etc… The possibilities are endless. Prosodic elements are classified as "suprasegmental" in that these informational units lie "above the segment", or above the immediate meaning of an utterance. Segmental units without attached prosody have the capacity to reveal communicable meaning, as is often captured by written language. However, suprasegmental elements are critical to the full understanding of meaning in language. While it is impossible to reliably capture prosodic force in writing, we often attempt to mirror the phenomenon using elements of punctuation, such as the comma to indicate a pause, or other stylistic adjustments, such as CAPITAL LETTERS to indicate heightened emotion. Much of the information that prosody supplies, however, cannot be captured through traditional orthographic means (Silverman 1992: 867). In a sense, the role of

prosody in language is supplementary at the same time that is necessary; it is essential for complete communicative competency, but it cannot exist without the original presence of segmental language. In sum, prosody comprises all vocalized aspects of speech that are separate from, though concurrent to, the individual phonetic units of language.

In approaching the role of prosody in semantics, it is helpful to appreciate the distinction between tonal and non-tonal languages. The prior presentation of prosody as an aspect of speech that is *supplementary* to basic meaning generally pertains to non-tonal languages, such as English. In contrast, tonal languages systematically use pitch to communicate and distinguish between the lexical or grammatical meaning of words. Tone is just as critical as the phonemic value of a word in determining its meaning (Pytlyk 2008:3). For example, in Mandarin, the morpheme /ma/ can be expressed using four different tonal patterns to communicate four different lexical meanings (See Example 1a). In these languages, tone, which is just one aspect of prosody, is critical to the primary semantics of every lexical unit; tones are contrastive features, meaning that two words with identical segmental features but varying tonal patterns form "minimal pairs". The English language occasionally utilizes variations in intonation to distinguish between lexemes, as can been seen with some heteronyms (See Example 1b); while most words in the English language have the capacity to carry different meanings in different contexts, instances of relying on varied intonational patterns for disambiguation purpose are rare (Rodd 2004: 89). Other languages, such as Arabic and Italian, can utilize vowel duration as a linguistic element to assign the specific meaning of individual words (See Example 1c). The role of prosody in different languages is, therefore, highly variable.

(1a) Use of tones to distinguish between /ma/ morphemes in Mandarin. (Pytlyk 2008:3)

(a) mā (high level) "mother"

(b) mǎ (fall-rise) “horse”

(c) má (rising) “hemp”

(d) mà (falling) “scold”

(1b) Use of tone to distinguish the homonym “present” in English (capitalization indicates syllabic stress)

(a) PREsent – “a gift”

(b) preSENT – “to demonstrate”

(1c) Use of vowel length to distinguish between /ban/ morphemes in Arabic.

(a) [ban] “coffee beans”

(b) [ba:n] “to appear”

In order to truly explore and discuss “prosody” beyond the mere descriptive, it is critical to understand the methodological approach that linguists use for coding, systematizing, and qualifying prosody. A prevalent theoretical framework that linguists use to label the phonological system of English prosody is ToBI (Tones and Break Indices); this system was originally conceived as a means to standardize the transcription of prosody in a manner analogous to the standard provided by the International Phonetic Alphabet for transcribing individual phonetic units (Silverman 1992:867). While the ToBI scheme was not used in the design and analysis of the current project, it does provide a helpful approach for understanding the prosodic system of English. The ToBI system splits the components of prosody into two major tiers: tones and break indices. The break index tier aims to annotate the degree of phonological association between adjacent words. The tonal tier aims to transcribe the unidirectional sequence of pitch events that occur in a speech utterance (Silverman 2002: 868). There are two types of tones in English: (1) pitch accents and (2) phrase boundaries, and they are

built from two possible tone levels, High (H) and Low (L). The two phrasal boundaries/tones in English are located at the edges of phrases and are demarcated by the symbol "%". The five pitch accents in English can consist of either one or two of these tones, and they reflect the various ways in which these two tones can be combined (See Example 2); these pitch accents also pinpoint the tone that is phonologically associated with the primary stress of a syllable using the symbol "*" (Jun 1996:28). The type and placement of specific pitch accents in English are critical to the marking of information structure of an utterance. Meanwhile, boundary tones in English signal the illocutionary force or intent of an utterance. However, the relationship between the form and function of these facets is arbitrary and must consequently be learned by the speaker (Chen 2009: 370).

(2) Pitch accents in English with typical linguistic function (adapted from Pitrelli 1994:123, Hirschberg 2004:5)

H* simple high (declarative statements, new discourse information)

L* simple low (yes-no question, salient information)

L+H* low-to-high rise (contrastive focus)

L*+H scooped rise (pragmatic uncertainty)

H+!H* fall to stress (pragmatic inference, implied familiarity)

An entity as complex as prosody, however, will not be adequately captured by a standard without first prompting a number of problems. In the original presentation of the ToBI framework, there was a significant degree of transcriber variability in classifying the various types of pitch accents, with consistency rates ranging from 60-80% (Silverman 2002:869). Further studies have revealed that even experienced transcribers can have difficulty distinguishing between some of the pitch accents, such as H* and L+H* (Pitrelli 1994:123). I set

out to examine the relationship between semantic focus, utterance stress, and pitch change, which is why I initially approached this project through the lens of the ToBI system. However, while academic measures have been taken to gradually smooth out variability in the transcription of English prosody, the time constraints and level of expertise necessary to effectively use this system were prohibitive for my particular project. Instead, I utilized the general concept of the pitch accent as marker of stress in the English language to guide my coding of speech samples.

To grasp my method, consider the following recording of a speech sample in the program *Praat*:

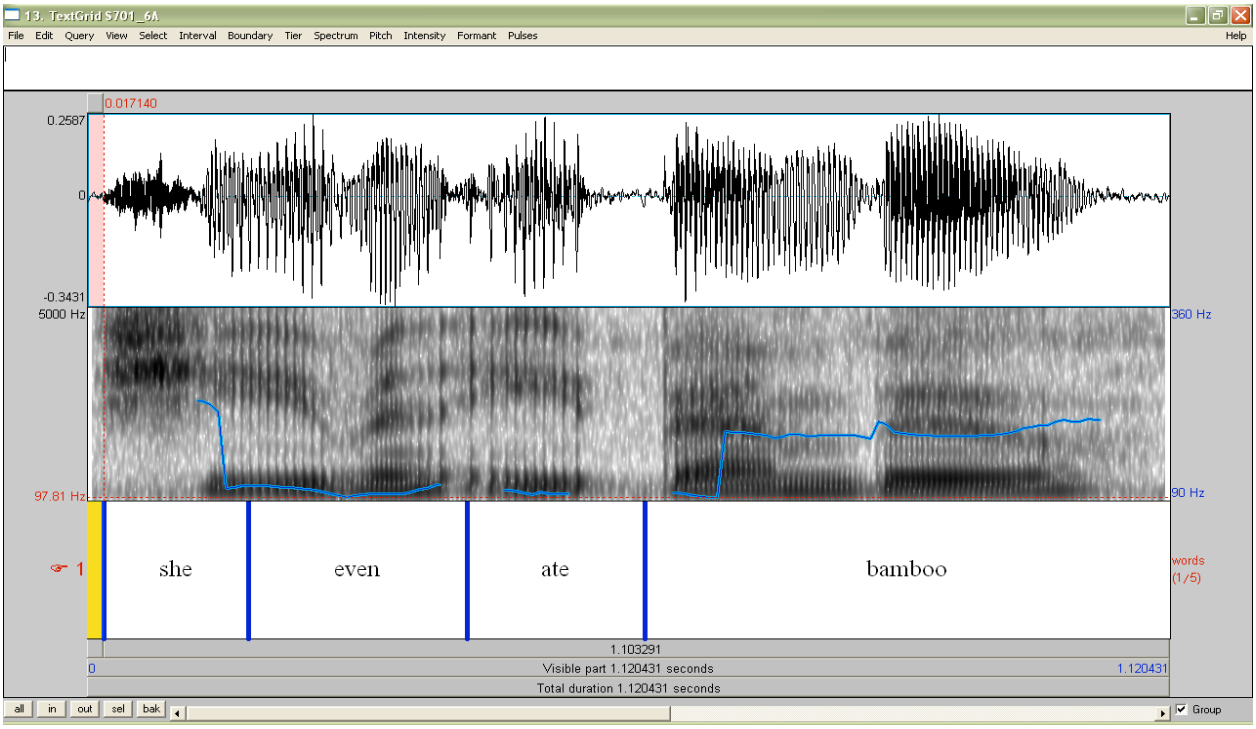

Figure I: Spectrogram of a recording of "She even ate bamboo"

In the context of the paragraph read by the speaker of this recording, the semantic focus of the utterance is the word "bamboo". From a purely qualitative overview of this recording, it is evident that the pitch level indicated by the blue line remains constant and low between the words "even" and "ate, but jumps to a higher pitch at the word "bamboo." This speaker is

marking the stressed focus of the sentence by placing a high pitch accent on "bamboo." Instead of demarcating every accent structure in the manner suggested by the ToBI system, I chose to focus on pitch changes between critical semantic points within the utterances. In this way, I attempted to capture the centers of stress and corresponding pitch accents within the recordings.

It is important to note that complete ToBI transcriptions would not have revealed much more information than I already extracted by examining pitch changes between specific points. In theory, the ToBI system captures a host of prosodic elements in its coding methodology, but the chief focus of the ToBI analysis still revolves around pitch. I mined comparable information from the data that ToBI would have yielded, but merely without engaging the vocabulary of the ToBI construct. I do not believe that *not* using ToBI detracted from the study. Perhaps as a result of my inexperience with the ToBI system, I also found the coding of pitches via purely qualitative means to be uncomfortably subjective. Instead of relying on auditory cues to determine the vocalized stress within a given recording, I employed quantitative measures to minimize any personal auditory bias as I coded for changes in pitch within the speech samples.

III.B. Prosodic transfer

Language systems vary quite significantly in their dependence on prosody as a means to convey meaning (Chen 2009: 367). Consequently, the extent to which an individual relies on intonational strategies for relaying messages is partially dependent on their native language. Because of these cross-linguistic differences, researchers have devised various contrastive categories to describe the general patterns of pitch reliance in a given language. For example, in noting that speakers of different languages have markedly different accents to indicate the focus of an utterance, Krahmer presents two different "types" of languages: plastic and non-plastic.

Languages that rely more heavily on word order for the purpose of marking information are often constrained in their ability to adapt their intonational structures; these are considered "non-plastic" and include languages within the Romance family. In contrast, "plastic" languages, such as English and Dutch, have a flexible word order that facilitates the molding of prosodic patterns to fit informational structure (Krahmer 2007:3). This varied reliance on structural and pragmatic rules can be understood along a qualitative scale of pitch accent patterns (See in-text Figure II).

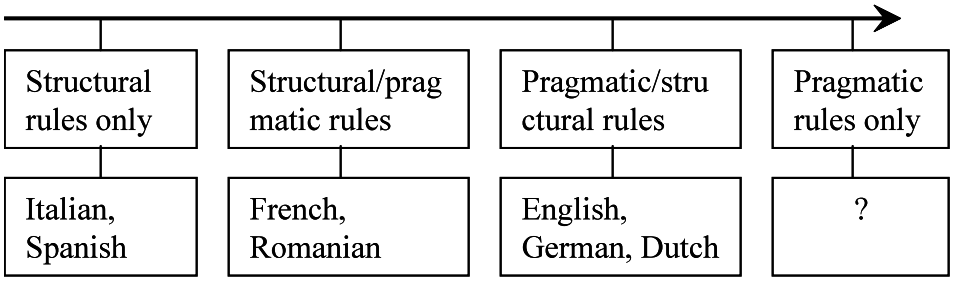

Figure II: Typology of accent systems. Adapted from Rasier 2007:53.

Given these complex differences in prosodic structure between languages, recent research has attempted to broach the question of how a given individual adapts to the prosodic system of a second language (Chen 2009, Krahmer 2007, Rasier 2007, Goad 2006). In brief, is it possible to adopt the pitch accents and phrasings of a second language (L2), or is that acquisition process inhibited by the prosodic system of a speaker's native language (L1)? And if the latter, what aspects of an individual's L1 affect their success in mastering the L2 prosodic structure? This area of study is broadly categorized under the concept of prosodic transfer.

Interest in cross-language interference in language learning arose during the 1950's in the form of the Contrastive Analysis Hypothesis (CAH). The original theory posits that in order to predict the areas of difficulty in second language learning, it is necessary to compare and contrast the structural elements of the native and target languages (Eckman 1977:315). The strong version

of the CAH predicts that similar features will be easy to acquire and dissimilar features will be difficult to acquire; the weaker version merely sees this model as a very general heuristic for second language acquisition (Eckman 1977:316-317). However, a later revision of this theory criticized the original concept for not taking into account the concept of directionality of difficulty. This revision, known as the Markedness Differential Hypothesis, suggests that the relative degree of difficulty in language learning can be determined using a theory of universal grammar that is independent of specific languages (Eckmann 1977:315). Markedness is defined using the following parameters: "A phenomenon 'A' is more marked than 'B' if the presence of 'A' implies the presence of 'B' but the presence of 'B' does not imply the presence of 'A'" (Eckmann 1977:320). For example, in all languages, the presence of voiced obstruents necessarily anticipates the presence of voiceless obstruents, but the converse is not true. Voiced obstruents would be "more marked" than voiceless obstruents, and it would therefore be more difficult to learn new voiced obstruents than new voiceless obstruents in a new language (Eckmann 1977:321). This hypothesis is supported by Menner's work with German and English learners in which native German speakers were much less capable of acquiring word-final voiced obstruents than native English speakers in the reverse direction (Eckmann 1977:317). While the previous example pertains to phonological transfer, the theory has been applied to prosodic transfer as well. Using a similar logic of markedness, Rasier asserts through his research with learners of Dutch and French that is should be easier for speakers of plastic languages to shift to the accentuation patterns of non-plastic languages than the converse (Rassier 2007:55). Supplementing these findings, research by Goad et al. has shown that features of L1 prosodic structures do not completely crystallize. Her work with Mandarin-to-English transfer revealed that it is even possible for speakers to partially adopt L2 prosodic features that are completely

absent in their L1 (Goad 2006:243). These observations all pertain to the present study in that the languages of interest, namely English, Korean, and Spanish, have quite distinct prosodic structures. Variation in broad descriptors such as "plasticity" among these three languages could provide the key points necessary to understand difficulties associated with prosodic transfer in second language acquisition.

Much of the literature has focused on the general interference of a speaker's L1 on their L2 that results in what is often perceived as a "foreign accent." These studies often limit their scope to the production and perception of individual phonemes in second languages, excluding an analysis of prosodic features. However, a recent study argues that L2 segmental errors are much less detrimental to listener comprehension than errors in prosodic phrasing and structure (Rasier 2007:43). Interestingly enough, areas of prosodic difficulty seem to be somewhat common across L2 language learners, regardless of the target language; these areas include: (1) manipulating L2 prosody in the correct contexts; and (2) conveying and perceiving linguistic meaning through the use of prosodic cues (Rasier 2007:44). Unfortunately, second language learners are rarely given instruction on the acquisition of intonational patterns in their non-native language (Chen 2009: 367). Because the effect of a speaker's L1 prosodic structure on their L2 prosodic development is not insignificant, the process of prosodic transfer is deserving of further study.

An early study by Delattre did, in fact, strive to codify the prosodic differences between Spanish and English speakers in order to predict difficulties in second language acquisition; however, his conclusions do not provide a broad enough foundation for the current study (Delattre 1963). A study by Adams that sought to define the acoustic correlates of "stress" within an utterance found that, in English, stress is often associated with greater changes in fundamental

frequency (larger pitch changes), greater duration (longer vowels), and greater falls in amplitude following a peak (loud followed by soft) (Adams 1978:144). However, native and non-native English speakers produced and perceived stress differentially. He posited that prosodic cues associated with conveying stress are developed in the framework of one's native language, which can interfere with production patterns in a speaker's second language (Adams 1978:154).

Of the more recent research that has been conducted on prosodic transfer to English, the studies have included native speakers of Korean and Spanish, both of which are groups that were recruited for my study. Ramírez Verdugo's work with native Spanish speakers suggests that non-native intonation patterns play a significant role in defining information structure and transmission in L2 English learners (Ramírez Verdugo 2006:26). She hypothesizes that differences in intended pragmatic communication stem from that fact that the nuclear pitch accent is not relied upon in Spanish as much as it is in English for defining the focus of an utterance. She suspects that a prohibitive barrier results from native and non-native differences in the phonetic realization of focus and direction (27). Spanish speakers do not rely on intonational structure as the predominant mechanism for ascribing discourse prominence. In fact, a study by Kelm showed that native English speakers learning Spanish tended to use increased intensity and higher pitch to indicate focus within their L2; in comparison, the native Spanish speakers used lexical and syntactic markers to identify contrast within the discourse (Kemp 1987:627). A similar study by Wennerstrom suggested that non-native English speakers of Spanish L1s did not use pitch accents to convey meaningful contrasts in a systematic manner (Wennerstrom 1994:399). Such observations are critical for understanding why native Spanish speakers might have difficulty engaging appropriate intonation as a primary linguistic resource in English as a second language (Ramírez Verdugo 2006:27).

In Korean the pitch accent serves a structurally distinct role from that found in English. The concept of the pitch accent in English is typically associated with syllabic stress within a word; as described by the English ToBI system, variations in pitch shape resulting from changes in accent prominence provide a means to contrast intonational meanings. In contrast, pitch accents in Korean serve to demarcate prosodic grouping of words (Jun 1996:30). Korean has neither lexical stress nor lexical pitch accents. Instead, the ebb and flow of pitch is linked to marking phrase location. This particular element of Korean prosody is termed a *maltomak*, which literally translates to "unit of speech" (Jun 2005: 203). However, while English and Korean differ in the shape of their tonal contours, focus is realized in both languages by raising the pitch range and increasing the amplitude of speech (Uleyama 1996:3). Due to these broader prosodic differences, it is perhaps not surprising that native Korean speakers also exhibit difficulty in the application of correct pitch accent prosody in English as their second language (Uleyama 1996: 18).

III.C. Types of Semantic Ambiguity

This sweeping discussion of prosody provides the background for attempting to envision prosodic tools as mechanisms for confronting semantic ambiguity. The intersection of these two concepts forms the core of this research. Before delving into the specifics of the current project, it would be helpful to hone in on the work of a few individuals who have tackled similar questions. Over a series of three studies, Julia Hirschberg and Cinzia Avesani compared and contrasted the various prosodic cues that speakers use in their native languages to disambiguate semantically ambiguous utterances. They sought to address two key points: (1) whether the

presupposition that intonation *is* used as a mechanism for disambiguation actually holds; and (2) if such patterns of disambiguation are consistent cross-linguistically (Hirschberg 1997:1).

The Hirschberg studies examined a number of different kinds of semantic ambiguity, including focus sensitive operators, scope of negation, and quantifier scope. Four of these types of ambiguity were used in the current study: the focus operators "only" and "even", the scope of negation construction "not-because", and the quantifier scope "every". Before proceeding further, it is critical to understand how each of these linguistic constructions causes semantic ambiguity, yielding incongruent logical meanings in different contexts.

III.C.i. "only"

As is perhaps evident by its label, focus sensitive operators are words that can take varying focus in an utterance. Focus operators direct the stress of a phrase by setting up contrastive information. The focus operator "only" is used to indicate the concept of singularity or uniqueness within an utterance. Consider the phrase "She only hit John." This utterance consists of a subject ("She"), a focus operator ("only"), a predicate ("hit"), and an object ("John"). In this phrase, the focus operator could hold focus over either the predicate or the object. However, without an external context or other linguistic cue, it is impossible to unambiguously determine whether the "focus" of the focus operator is the object or the predicate. Now consider this phrase in the following two contexts:

(3a) FOCUS=OBJECT

"Anne could have hit multiple people that day, but **she only hit John**."

(3b) FOCUS=PREDICATE

"John has been such a bad boyfriend to Anne. When they broke up, I'm surprised **she only hit John**."

In example (3a), it is very clear that the focus operator is setting up contrastive information with the object "John." The actor hit "John" as opposed to hitting "Sarah", "Gene", "the cat", "the mattress", or any other object. The focus of the operator is, therefore, the object. In example (3b), the context again makes the meaning of the utterance clear. In this scenario, the object "John" is given information; it has already been introduced in the discourse. The focus operator serves to clarify that the actor was "hitting" John as opposed to "kicking", "scratching", or "biting" him. The focus of the operator is, therefore, the predicate "hit." It is in this manner that we can identify the two potential focuses of the focus operator. Using predicate logic via formal semantics, the resulting interpretations are as follows:

(4a) FOCUS=OBJECT

HIT(anne,john) & **(**∀x (PERSON(x) & x≠john)→ ~HIT(anna,x)))

"Anne hit John, and for all x, if x is a person and x is not John, then Anne did not hit x."

(4b) FOCUS=PREDICATE

HIT(anne,john) & **(**∀P λx.λy. P≠HIT(x,y))→ ~P(anne,john)))

"Anne hit John, and for all P, such that P is a predicate relating entity x to entity y, if P is not hit, then there is no predicate that relates Anne to John."

III.C.ii. "even"

The focus operator "even" serves to indicate an action, event, or object that extends beyond the normal range of expectation within a discourse. It is very similar to the focus operator "only" in that it can also take an object and a predicate focus. Consider the two contexts below:

(5a) FOCUS=OBJECT

"Jacob is a compulsive pyromaniac. Last week, **he even burned a cat**."

(5b) FOCUS=PREDICATE

"Jacob has a strange affinity for hurting cats. He often kicks them, pulls out their hair, and pours water on them. Last week, **he even burned a cat**."

As was seen with the "only" focus operator, the meaning of the semantically ambiguous utterances is very clear within each of these two contexts. In example (5a), the concept of "burning" is given information. It is understood that "Jacob" has "burned" items in the past. In contrast, the speaker indicates that the burning of the "cat," as opposed to any other object, is the central stress of the utterance. The focus of the focus operator is, therefore, the object. Utterance (5b) provides an example in which the object "cat" is old/given information. The actor has already conducted a number of actions towards "cats." The focus operator indicates that the selected predicate, "burned", is a particularly notable among the various actions that "Jacob" has taken against the object "cats". Therefore, the focus of the focus operator is the predicate.

The ambiguity derived from unresolved scopes of negation and quantity is quite different from that found within utterances containing focus sensitive operators. In phrases with ambiguous focus operators, the denotative meaning of the two competing contexts remains essentially identical. Consider again the utterance "Anne only hit John." Regardless of the intended focus of the operator, the phrase still indicates that some individual named "Anne" did, in fact, perform the action of "hitting" some individual named "John" at some time point in the past. The resulting denotation of the statement can be logically described as follows:

(6) HIT(anne, john)

"Anne hit John."

What varies is that focus operators allow speakers to emphasize and accentuate the intended stress of an utterance, which, as has been shown, can vary between contexts.

III.C.iii. “not-because”

With ambiguity that arises from disparate scopes of negation, the two competing interpretations yield contrastive denotations. The easiest way to grasp this idea is through another set of example contexts. Consider the semantically ambiguous phrase “Joseph isn’t drinking because he’s unhappy.”

(7a) BECAUSE>NOT (narrow scope of negation)

> “Joseph usually loves to go out drinking. Tonight, however, **Joseph isn’t drinking because he’s unhappy.** We should talk to him.”

(7b) NOT>BECAUSE (wide scope of negation)

> “Sometimes people drink when they’re depressed, but that isn’t the case tonight. **Joseph isn’t drinking because he’s unhappy.** He’s drinking because he just got a promotion.”

This example poignantly captures the rather strange notion that one single utterance can have two logical interpretations that have meanings with contradictory conclusions. In the narrow scope of negation example (7a), it is clear from the context that “Joseph” is both “unhappy” and “not drinking”; in fact, his “unhappiness” is the direct cause of his “not drinking.” The scope of negation falls on the predicate, and the subordinate clause serves to provide an explanation for why “Joseph is not drinking.” Because Joseph is unhappy, he does not drink.

In contrast, the utterance in example (7b) yields quite a different meaning. In this context, “Joseph” is both “drinking” and “not unhappy”. The scope of negation in this scenario falls over the explanation. The purpose of the word “isn’t” is to negate the reasoning that is presented by the subordinate clause beginning with “because.” Unhappiness *is not* the reason for Joseph’s drinking. One key point to note, however, is that we cannot definitively conclude the state of Joseph’s happiness in the wide scope of negation. We can conclude that Joseph is drinking, and the context *implies* that Joseph is, in fact, happy. However, as with most instances of pragmatic implicature, it is possible to negate that intimation. Consider example (7c):

(7c) NOT>BECAUSE (wide scope of negation)

> "Joseph is really unhappy. However, that's not the reason for his behavior today. **Joseph isn't drinking because he's unhappy.** He's drinking because there's an open bar at the party."

Thus, scope of negation ambiguity with "because" clauses is quite complex. The two main interpretations, however, are as follows:

(8a) BECAUSE>NOT

> "Joseph is unhappy and Joseph is not drinking. In fact, his unhappiness is causing him to not drink."

(8b) NOT>BECAUSE

> "Joseph is drinking, and unhappiness is not the reason for why he is drinking."

III.C.iv. "every"

The fourth type of ambiguity that I explored involves dueling scopes between quantifiers and negation. Quantifier ambiguity is, perhaps, the most abstract type of ambiguity examined in this project. As with the "not-because" ambiguity, the two interpretations that result from these semantically ambiguous phrases have quite distinct meanings. For example, consider the following two contexts of the phrase "every lizard didn't jump over the ramp":

(9a) EVERY>NOT (quantifier scopes over negation)

> "Three lizards were riding around on motorcycles. They were trying to jump over a ramp. None of them made it. **Every lizard didn't jump over a ramp**."

(9b) NOT>EVERY (negation scopes of quantifier)

> "Three lizard were riding around on motorcycles. They were trying to jump over a ramp. The green and orange lizards were able to jump over the ramp. However, the black lizard was unable to jump over the ramp. **Every lizard didn't jump over a ramp."**

In example (9a), it is clear that no lizard jumped over the ramp. There was a certain activity, namely "jumping over a ramp", that all individuals failed to complete. None of the lizards jumped over the ramp. In a sense, the quantifier phrase has priority over the negation. In context

(9b), the interpretation of the phrase is conspicuously different. As informed by the sentences preceding the phrase, there were some lizards that did jump over the ramp, and there was one lizard that did not jump over the ramp. The context of the phrase indicates that at least one lizard failed to jump over the ramp. This situation might often be glossed as "not every lizard jumped over the ramp." The quantifier is not all-encompassing. Therefore, the negation scopes over the quantifier. In formal semantics, this distinction can be clarified using predicate logic statements, as seen in example 10:

(10a) EVERY>NOT

∀x (LIZARD (x) → ~JUMP_OVER(x,ramp))

For all "x", if "x" is a lizard, then "x" did not jump over the ramp.

(10b) NOT>EVERY

∃x (LIZARD(x) & ~JUMP_OVER(x, ramp))

There is at least one "x" such that "x" is a lizard and "x" did not jump over the ramp.

A schematic of the four types of ambiguity is shown in Figure III.

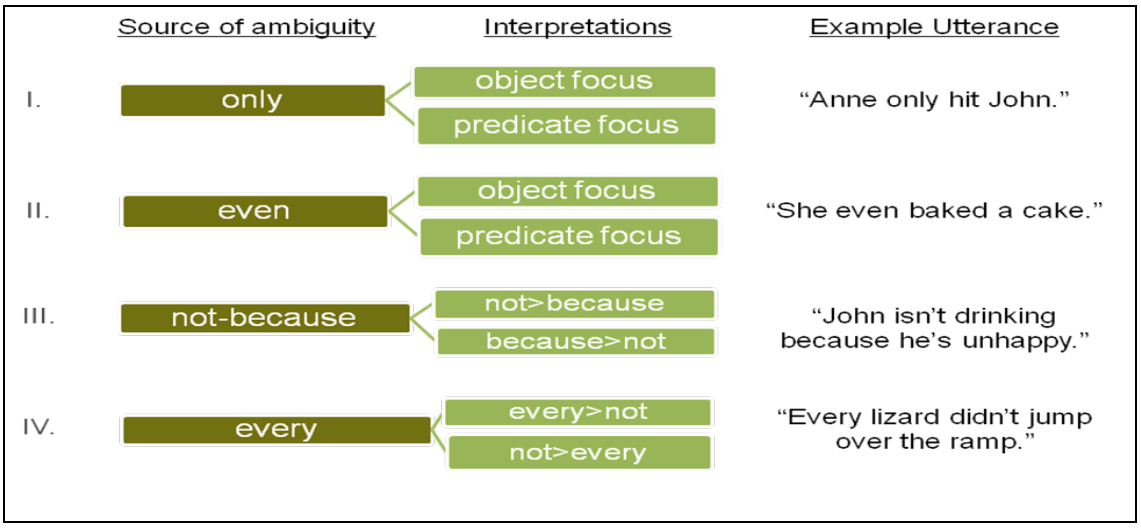

Figure III: Classes of semantic ambiguity with corresponding scopes and focuses used in this study

III.D. Prosody & Semantic ambiguity – Past research

As mentioned previously, Julia Hirschberg and her colleagues examined prosody as a mechanism in speech production for resolving these types of semantic ambiguity in native speakers of English, Spanish, and Italian. In the utterances involving focus sensitive operators, they found that English speakers tended to realize the nuclear "stress" of an utterance on semantic focus of the operator in 64/72 trials (Hirschberg 2000:90). Italian speakers disambiguated these phrases less frequently, but when they did, they placed a high pitch accent on both the focus sensitive operator and the focus itself (Hirschberg 1997: 3). Spanish speakers were inconsistent with their patterns of focus sensitive disambiguation. A related perception study by Nambu & Lee suggests that the focus operator *ocik* "only" in Korean requires a salient prosodic marker to develop an association with the focus element of the utterance. Participants classified utterances that exhibited higher pitch on the focus of *ocik* as the most natural (Nambu & Lee 2010:2). However, this finding is rather curious since, according to Jun, pitch accents are not typically used in Korean to indicate the stress of an utterance (Jun 1996:30).

In the Hirschberg studies, the clearest patterns of prosodic disambiguation occurred with the scope of negation. In all three language groups, speakers articulated the wide scope of negation (not>because; see Example 7b) as one single intermediate phrase, and the narrow scope of negation (because>not; see Example 7a) with two distinct phrases marked by a phrase boundary (Hirschberg 2000:87). However, speakers of Spanish, Italian and English further disambiguated these contexts using different prosodic mechanisms. Using ToBI contours, Hirschberg claimed that in 90% of cases, English speakers ended the wide scope utterances with a High boundary tone and the narrow scope utterances with a Low boundary tone (Avesani 1995:2). In contrast, Italian and Spanish speakers used pitch accents to mark the phrase boundary

in the narrow scope of negation; in wide scopes of negation, Italian and Spanish speakers placed a pitch accent on the predicate of the main clause and then de-accented the lexical item in the subordinate "because" clause (Avesani 1995:3). A comparable production study by Baltazani showed that native Greek speakers also produce two consistent patterns to link prosody to information structure in these not-because contexts (Baltazani 2002:160).

With ambiguity involving the quantifier "none", Hirschberg found no clear patterns of prosodic disambiguation within English speakers (Hirschberg 2000:91). There were some patterns of disambiguation within Spanish and Italian speakers, but these strategies were inconsistent and varied among individuals. As a general observation, however, the prosodic patterns of Italian and Spanish speakers were more similar to each other than to those of the native English speakers (Hirschberg 1997:4). Baltazani also examined a type of quantifier ambiguity in which two different quantifiers are competing for scope. She found that Greek speakers systematically produced two different patterns of disambiguation, but did not distinguish between patterns in perceptual analysis (Baltazani 2002: 166). However, neither of these two studies looked specifically at the relationship between negation and quantifier scope.

III.E. Changes in the current study

While these studies have opened the line of inquiry regarding the linguistic intersection of prosody and semantic ambiguity, a number of questions remained unresolved. For the current study, I used Hirschberg's design as the basis for constructing the paradigm of the experiment. However, I sought to build on her work by approaching the question from a number of different angles. In addition to examining the speech contours of native English speakers, I also examined the speech of non-native English speakers with native languages of distinctly different prosodic

structures. The relative ability of English L2s to acquire and match the prosodic patterns of English L1s in these semantically ambiguous contexts could actually provide meaningful information as to prosodic function in native speakers. Furthermore, these L2 results should determine if prosodic assimilation is necessary for comprehension of these semantically ambiguous utterances.

As previously mentioned, I also amended the experimental design by including a perception study. Since prosody is only relevant to language as a communicative tool if it is both encoded by the speaker and comprehended by the listener, I believe that it is critical to examine the full speech circuit (Krahmer 2007:1). Furthermore, the use of a paradigm with identical semantic contexts in both the production and perception portions of the study should readily allow for a comparison of the results. I hoped to establish an intricate link between the two studies by using excised speech recordings from the production portion in the construction of the perception portion.

In the discussion of her design, Hirschberg noted that the experimental paragraph contexts might have made the semantic interpretations too obvious (Hirschberg 2000:94). A feature of prosody that I have emphasized heavily in this paper is that its role is quite malleable. Speakers have multiple linguistic tools at their disposal to signify intended meaning. Intonational prosody has the capacity to serve a host of functions, but it is important to note that inter-speaker and inter-language reliance on any given tool is highly variable. Different speakers employ different strategies to convey their messages (Hirschberg 2004:12). If speakers have the option to rely on the surrounding discourse of a semantically ambiguous utterance to fully resolve its meaning, they may not rely on prosodic means at all. With that criticism in mind, I attempted to maintain a minimalist mindset in the construction of the experimental paragraphs. Another

criticism that I personally noted in Hirschberg's study was that some of her example paragraphs were, albeit elaborate, quite difficult to comprehend. An oft expressed concern regarding production studies is that the linguistic goals of the speaker in reading a paragraph are quite distinct from those of engaging in a live conversation; therefore, the manner in which a person adjusts their intonation in an experimental setting might differ dramatically from how they sound naturally. However, if the researcher solely captured unprompted discourse, it would be virtually impossible to control for the target linguistic utterances (Schafer 2000: 169). Noting these observations, I attempted to find the balance between clarity and redundancy in designing the experimental contexts. I also incorporated mechanisms in the procedure of the production study to ensure that participants both understood the contexts and remained engaged with the study.

Finally, in my approach of the results, especially those of the production study, I attempted to be consistent and methodical. Prosody is a linguistic concept that encompasses a broad, almost overwhelming, scope of linguistic topics. In order to finitely address the question at hand, I chose to specifically examine pitch change as a prosodic marker in the speech samples of the production study. For the four types of ambiguity that I set out to explore, especially with the quantifier and negation scopes, there was not much precedent in the literature for determining the appropriate pitch markers to examine. In a sense, I delved into the coding and analysis of these pitch differences without having a full grasp of what prosodic patterns to expect. However, I compensated for this lack of certainty by being consistent with my coding. Whether or not this consistency yielded significant results in all cases, I wanted future researchers to feel comfortable relying on my results in determining which prosodic markers to further study. Additionally, I made the decision to code my production results using a system of relative pitch changes distinct from that of ToBI contours. I realized at an early point in my project that I

would not have the opportunity to receive training in the ToBI system to a degree that I would feel comfortable accurately coding my speech samples. Given that I chose to examine pitch changes between specific points in these utterances, however, I do not believe that a full ToBI analysis would have yielded much more useful information.

In the next sections I will provide a thorough outline of my experimental design and results. Before proceeding, however, I would like to re-orient the paper with respect to the major guiding questions of the project. First, I hope to study the role of one particular mechanism, prosody, in the resolution of ambiguity that is inherent to the linguistic system. Second, I hope to explore the concept of prosodic transfer; does the prosodic structure of a speaker's native language interfere with their ability to adapt to a novel prosodic system? And finally, do we perceive the same patterns that we produce in our interpretations of semantic ambiguity?

II. Methods

The study was "Exempt Approved" by the Emory University Institutional Review Board (IRB). For reference, the study number is IRB00044462.

For the purposes of implementing the production and perception portions of this study, I devised a series of forty (40) paragraphs that corresponded to twenty (20) semantically ambiguous utterances. Each semantically ambiguous sentence had two corresponding paragraphs that disambiguated the meaning of the utterance through context. For example, with the utterance "Mrs. Cohen only baked the cake", the two corresponding paragraphs were as follows (See Appendix B for the full list of contexts):

(11a) BECAUSE>NOT (narrow scope of negation)

Harold really dislikes his literature class, and after his four-hour swimming practice today, he just doesn't have the energy to deal with Herman Melville. **Harold isn't reading because he's tired.**

(11b) NOT>BECAUSE (wide scope of negation)

Harold isn't reading because he's tired. Harold is reading right now because he has a midterm tomorrow.

Two separate Microsoft PowerPoint presentations were designed, each with 20 of these paragraphs. Only one of the contexts for each semantically ambiguous utterance was placed within a given PowerPoint presentation. The two presentations were blocked in such a way as to equally balance the total number of each of the four kinds of ambiguous utterances and the two semantic interpretations present within each of those types of ambiguity. Each paragraph was placed on a separate slide. The order of the twenty paragraphs was selected randomly. Following each of these paragraphs, I placed a forced-choice question to determine if the participants had interpreted the semantic context correctly. For example, following the paragraphs from Example 11, I posed the question "Is Harold reading?" The purpose of placing these questions was also to ensure that the participants would remain actively engaged in the study.

For the production portion of this study, 24 undergraduate students at Emory University were recruited consisting of 8 native English speakers, 8 Korean-L1 English speakers, and 8 Spanish-L1 English speakers. Participants were primarily recruited via a "snowball effect" through online recruitment and word-of-mouth. The study was conducted in the Emory University Linguistics Transcription Lab. Upon arriving at the site, participants were assigned a unique identification number for coding purposes. They were then briefed on the study and asked to review and sign an informed consent form. They also filled out a brief survey documenting their personal language history.

The participants were recorded using an Audio-Technica AT2020 USB microphone and the software Audacity. Prior to the beginning of the formal experiment, participants were asked to speak a nonsense sentence in order to gauge if the microphone was capturing an adequate

level of sound for subsequent analysis. Participants were instructed to mentally read through the first context paragraph. It is important to note that the semantically ambiguous utterance within each paragraph was *not* accentuated by underlining, italics, or other stylistic means. It was not my intent for the participants to be overtly directed towards the source of potential ambiguity during their readings. Participants were then shown the forced-choice question that indirectly addressed the semantically ambiguous portion of the context paragraph. They were asked to nod their heads when they had finished processing the written contexts. They were *then* told to read the paragraph aloud and subsequently answer the yes-no, forced answer question. They then proceeded to the remaining paragraphs. Participants usually completed the recordings in about 10-15 minutes. Upon completing the study, participants were monetarily compensated for their involvement. Funds were generously provided by Scholarly Inquiry and Research at Emory through an Independent Research Grant.

For the perception study, I recruited 25 more undergraduate students from Emory University, consisting of 9 native English speakers, 8 Korean-L1 English speakers, and 8 Spanish-L1 English speakers. The reason for having a ninth native English speaker resulted from a communication error regarding the status of one participant's native language. This portion of the study was also conducted in the Emory University Linguistics Transcription Lab. The same informed consent, debriefing, and compensation procedures were used as in the production study.

For the perception study, I used a similar research design with a few functional tweaks. I created two more PowerPoint presentation and utilized the same semantically ambiguous utterances and paragraphs; however, I adjusted the manner in which the utterances were presented. The participants alternated between two different types of tasks. In the "speech

selection" task, participants were presented with one of the context paragraphs. They were then told to listen to two speech samples of just the semantically ambiguous portion of the contexts and select the one that they believed to correspond with the presented paragraph. The speech clips were excised from the recordings of the production study participants. By using recordings of individuals that had already read those paragraphs, I was able to control for the corresponding semantic target of the speech clips. In the "paragraph selection" task, participants were presented with the two context paragraphs that contained the same identical semantically ambiguous construction. They were also presented with one speech sample and instructed to select the paragraph that they believed to correspond to the recording. Each Presentation contained ten "speech selection" tasks and ten "paragraph selection" tasks. As with the production study, the presentations were balanced with equal numbers of each type of semantically ambiguous utterance and semantic focus/scope within each of those types.

At the beginning of the perception study, participants were given Sennheiser HD 201 headphones in order to minimize interference from external sources of sound. For each question, the participants were allowed to listen to the speech samples as many times as they desired. They were told to verbally express their paragraph and speech selections to the researcher as they progressed through the study. Although participants had control of their pace, the sessions never lasted longer than 15 minutes.

III. Results

V.A. Description of Production Analysis

After finishing the recording portion of the production study, I parsed out and labeled the specific semantically ambiguous phrase from each of the twenty paragraphs spoken by the

participants. Using the computer software *Praat,* I analyzed relative changes in pitch between certain elements within each of the four types of ambiguous phrases. In the utterances involving the focus operators "only" and "even", I examined: (1) the pitch change between the stressed vowel of the focus operator and the stressed vowel of the subsequent predicate (See Example 12a); and (2) the pitch change between the stressed vowel in the predicate and the stressed vowel of the subsequent object (See Example 12b).

(12a) Pitch change between the focus operator and the predicate.

He	even	burned	a cat.
	focus operator	predicate	object

(12b) Pitch change between the predicate and the object.

He	even	burned	a cat.
	focus operator	predicate	object

In this way, the pitch level changes of both potential arguments of the focus operators were examined. As described earlier in the paper, previous research suggests that a high pitch accent is often realized on the intended semantic focus of the utterance in English (Ramírez-Verdugo 2006:27). Coming from these experimental observations, I hypothesized that an increase in pitch level would be observed on the lexical element that corresponded to the semantic focus of a given utterance. Specifically, speakers should produce a pitch increase on the predicate for predicate-focus utterances and on the object for object-focus utterances. For example, in the following contexts I expected to see a pitch increase at the points indicated by the arrows:

(13a) FOCUS=PREDICATE

He	even	burned	a cat.
	focus operator	predicate	object

(13b) FOCUS=OBJECT

He	even	burned	a cat.
	focus operator	predicate	object

With the ambiguous phrases involving the scope of negation, I examined a number of different points within the sentences of interest. I broke down the changes in pitch levels between: (1) the word indicating negation, such as "isn't" or "didn't", and the subsequent predicate (See Example 14a); (2) the predicate and the beginning of the dependent clause, "because" (See Example 14b); and (3) the predicate and the stressed focus of the dependent clause (See Example 14c).

(14a) Pitch change between negation and predicate

Joseph	isn't	drinking	because	he's unhappy.
	negation	predicate	"because"	"cause"

(14b) Pitch change between the predicate and the "because" of the subordinate clause

Joseph	isn't	drinking	because	he's unhappy.
	negation	predicate	"because"	"cause"

(14c) Pitch change between the predicate and the stressed focus of the subordinate clause

Joseph	isn't	drinking	because	he's unhappy.
	negation	predicate	"because"	"cause"

Hirschberg restricted her analysis of scope of negation contexts examining "major and minor prosodic phrase boundaries" (Hirschberg 2000:9). However, on initial examination of the recordings, I was unable to identify noticeable, systematic differences in phrasing between the two semantic interpretations of the utterance. I could sometimes "hear" brief pauses and pitch boundaries between different phrases, but I was unable to visualize those features in *Praat*. For that reason, I attempted to focus my analysis of intonation on measurable pitch differences within the phrases. Based on my native speaker intuitions, I hypothesized that there would be a general pitch increase at the end of the utterance in the wide scope utterances, and a pitch decrease at the end of the utterance in the narrow scope utterances. As illustrated in schematic 15, I expected the pitch changes to occur at the points indicated by the arrows:

(15a) BECAUSE>NOT (narrow scope of negation)

Joseph	isn't	drinking	because	he's unhappy.
	negation	predicate	"because"	"cause"

(15b) NOT>BECAUSE (wide scope of negation)

Joseph	isn't	drinking	because	he's unhappy.
	negation	predicate	"because"	"cause"

In the phrases involving the ambiguous scope quantifier, I examined pitch changes between: (1) the stressed syllable of the word containing the scope quantifier and the word indicating negation (See Example 16a); and (2) the word of negation and the stressed focus of the remainder of the argument (See Example 16b).

(16a) Pitch change between the quantifier and negation

Every	lizard	didn't	jump over the ramp.
Quantifier		Negation	Argument

(16b) Pitch change between negation and the stressed focus of the remainder of the utterance

Every	lizard	didn't	jump over the ramp.
Quantifier		Negation	Argument

Hirschberg's research did not reveal any systematic prosodic differences between the two interpretations of the quantifier ambiguity utterances (Hirschberg 2000:91). Due to the lack of precedent, I did not have a clear intuition regarding the expected pitch changes in the two quantifier interpretations.

In addition to noting whether pitch had increased, decreased, or stayed level between any two points of interest, I classified the intonational change as "large", "medium", or "small" based on the relative frequency change in Hertz. Changes less than 5 Hz were classified as "level", between 5 and 25 Hz as "small", between 25 and 50 Hz as "medium", and greater than 50 Hz as "large."

V.A.i. “only”

For the focus operator “only”, 96 total recordings were collected and analyzed. The focus of these utterances was split evenly between the object and the predicate. To refresh, the semantic difference between the two contexts can be understood through example 3, replicated below. Prosodically, I expected speakers to register an increased pitch accent on the object in object-focus contexts and on the predicate in predicate-focus contexts.

(3a) FOCUS=OBJECT

“Anne could have hit multiple people that day, but **she only hit John**.”

(3b) FOCUS=PREDICATE

“John has been such a bad boyfriend to Anne. When they broke up, I’m surprised **she only hit John**.”

For each of the two semantic interpretations, there were sixteen representative recordings from each of three native language groups (See in-text Figures VI, VII; Appendix Tables 1, 2).

In the object focus utterances, the native English speakers registered a pitch decrease on the predicate 13/16 times and a pitch increase on the object 10/16 times. Furthermore, half of the recordings that showed such an increase on the object were classified as “large” changes, exceeding 50 Hz. Nine of the 16 recordings followed the sequential pattern of a pitch decrease on the predicate followed by a pitch increase on the object. Figure IV illustrates a spectrogram from the software *Praat* recorded by a native English speaker in the object focus context. This speech sample shows the expected pitch rise from the predicate “hit” to the object “Joseph”, as indicated by the blue pitch contour.

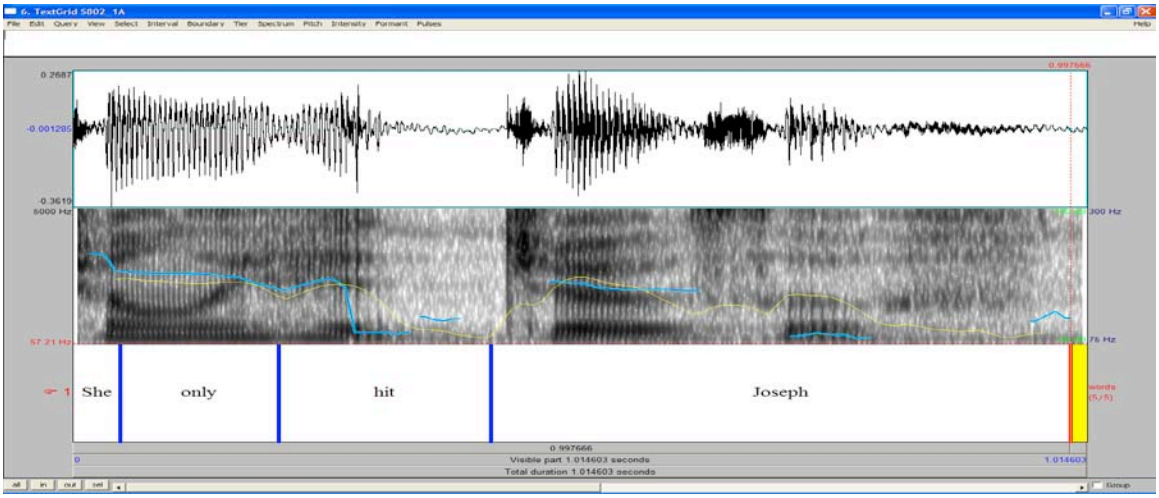

Figure IV: Spectrogram of object focus recording of "She only hit Joseph."

In contrast, in the predicate focus utterances, the native English speakers registered a pitch increase on the predicate 7/16 times and a pitch decrease on the object 11/16 times. Over half (4/7) of the utterances that showed a pitch increase on the predicate were also classified as "large" changes. The most common sequential pattern, 7/16 utterances, involved a pitch increase on the predicate followed by a pitch fall on the object. Figure V illustrates a spectrogram of a native English speaker recording in the predicate focus context. This sample reflects the hypothesized intonational pattern for the predicate focus context with increased pitch accent on the predicate "hit."

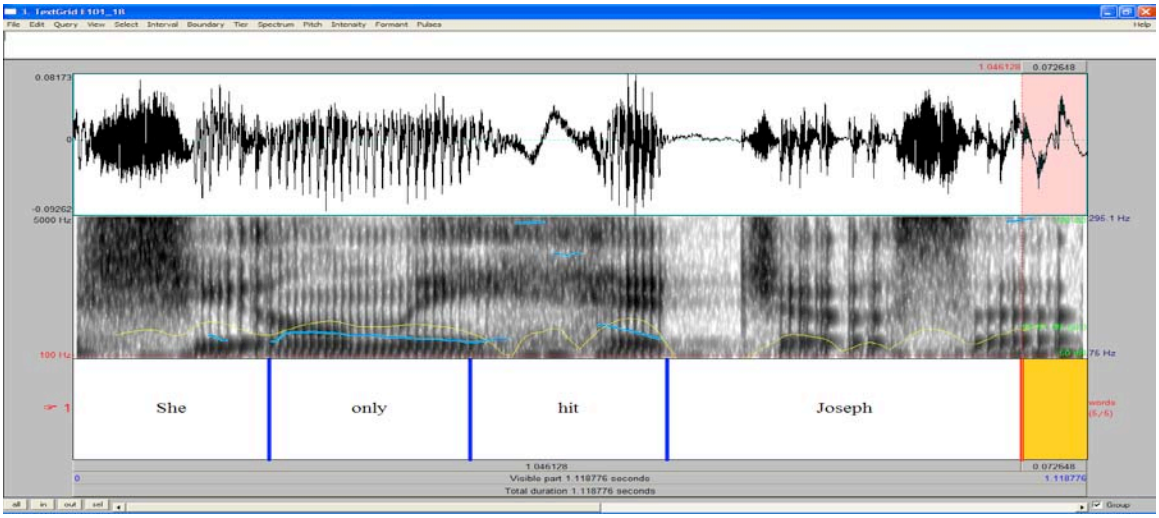

Figure V: Spectrogram of predicate focus reading of "She only hit Joseph."

For the native Korean speakers, the differences were not as distinct. In the object focus interpretation, 6/16 recordings registered a pitch increase on the object and 9/16 recordings registered a pitch decrease on the predicate. However, in looking at the combined pitch shifts, the most frequent pattern was a tie between complementary combinations: a pitch increase followed by a decrease (5/16) and a pitch decrease followed by an increase (5/16). In the predicate focus interpretation, 3/16 recordings yielded a pitch increase on the predicate and 10/16 recordings yielded a pitch decrease on the object. In fact, more Korean speakers produced a pitch increase on the object (5/16) than the predicate (3/16) in the predicate focus utterances. As with the object focus utterances, there was no clear intonational pattern for disambiguating the predicate focus utterances when the combinations were analyzed together.

For the native Spanish speakers in the object focus interpretation, 6/16 recordings registered a pitch increase on the object and 11/16 recordings registered a pitch decrease on the predicate. However, due to a large number of "level" pitch changes, only 1/16 recordings showed a pitch increase on the predicate, which makes the increase on the object more noteworthy. In the predicate focus utterances, only 3/16 recordings showed a pitch increase on

the predicate, but 12/16 utterances showed a decrease on the object. The most common pattern for the disambiguation of the predicate focus (5/16) was a decrease on the predicate followed by a subsequent decrease on the object.

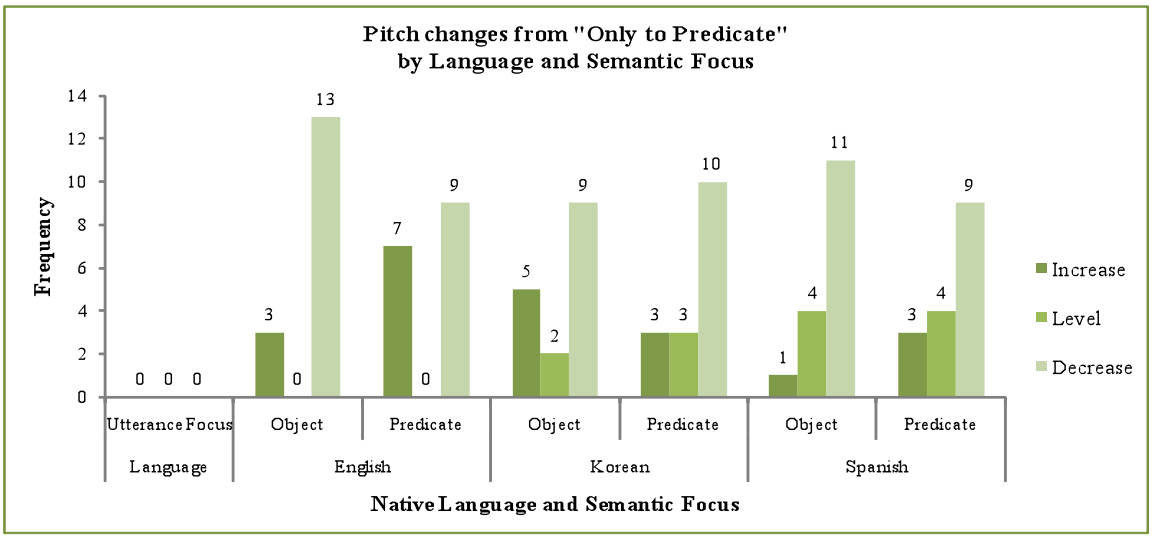

Figure VI: ONLY - Pitch changes from “Only to Predicate” by Language and Semantic Focus

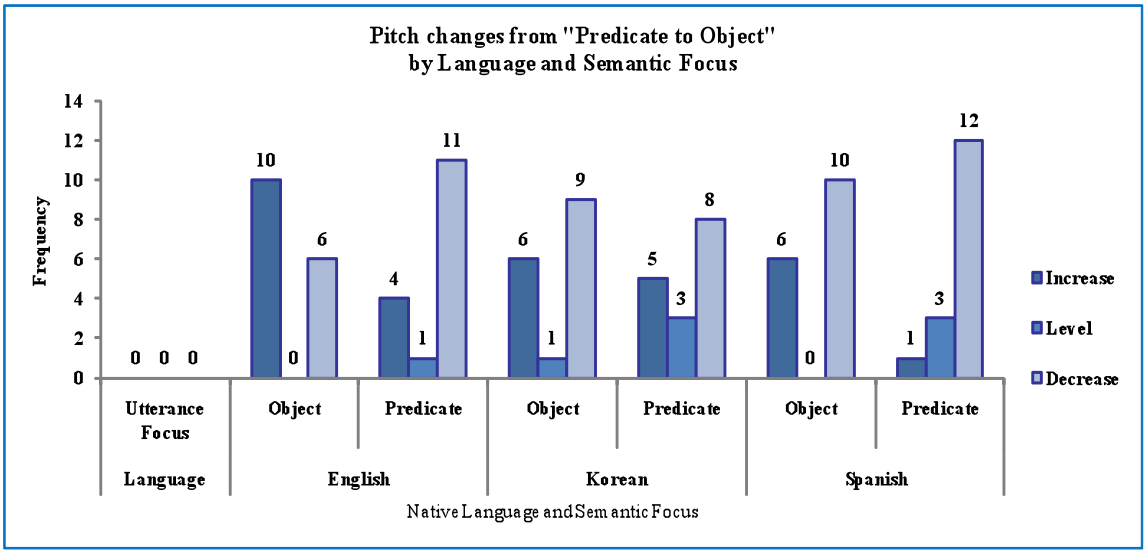

Figure VII: ONLY - Pitch changes from “Predicate to Object” by Language and Semantic Focus

To determine the significance of the frequencies of pitch increases and decreases between and within the three categories of speakers, I conducted a series of Fischer Exact Probability Tests using 2x2 Contingency Tables (See Appendix Table 3). Virtually none of the frequency differences were statistically significant at $p<0.05$ when analyzed with the more rigorous 2-tailed test. Nevertheless, I have provided a thorough presentation of the pitch change frequencies in order to show that some prosodic differences do seem apparent between native and non-native English speakers in the disambiguation of the focus operator "only." For example, within the native English group, the difference in the frequency of pitch changes between object- and predicate- focus utterances from the operator "only" to the predicate were not significant at $p=0.252$. However, the difference in pitch-change frequency from the predicate to the object held a significance value of $p=0.0731$. These numbers indicate that native English speakers are using different pitch patterns to resolve the semantic ambiguity that arises with the operator "only"; however, these frequency differences do not carry the statistical robusticity to generalize the pattern beyond the sample. The lack of statistical difference between some of these values still yields helpful information. For example, in native Korean speakers, the significance of pitch change differences between the object- and predicate- focus contexts was $p=0.676$ from "only" to the predicate, and $p=1$ from the predicate to the object. These numbers suggest that English-learners of Korean L1s are not utilizing consistent prosodic patterns at all to disambiguate between these two contexts.

V.A.ii. "even"

For the focus operator "even", 96 total recordings were also collected and analyzed. These utterances were split between object and predicate focus. The semantic difference between

the two contexts can be understood through example 4, replicated below. As with the focus operator "only", I hypothesized that speakers would register an increased pitch accent on the object in object focus utterances and on the predicate in predicate focus utterances.

(4a) FOCUS=PREDICATE

"Jacob is a compulsive pyromaniac. Last week, **he even burned a cat**."

(4b) FOCUS=OBJECT

"Jacob has a strange affinity for hurting cats. He often kicks them, pulls out their hair, and pouring water on them. Last week, **he even burned a cat**."

For each of the two semantic interpretations, there were sixteen representative recordings from each of the three native language groups (See in-text Figures VIII, IX; Appendix Tables 4, 5 for reference). In the object focus utterances, native English speakers produced a pitch fall on the predicate in 14/16 instances and a pitch rise on the object in 10/16 recordings. None of these speakers registered a pitch increase on the predicate. When analyzed in full, 9/16 utterances followed the dual pattern of a pitch decrease on the predicate with a pitch increase on the object. The patterns for the predicate focus utterances were less compelling, but still noteworthy in relation to the object focus contours. Only 5/16 recordings measured a pitch increase on the predicate, but 12/16 utterances measured a pitch fall on the object. Furthermore, half of those pitch falls were coded as "large", exceeding a decrease of 50 Hz. The most common combined pattern involved a pitch decrease on the both the predicate and the object (6/16).

In the object focus utterances of Korean speakers, only 4/12 recordings showed a pitch increase on the object. However, to put this finding in perspective, none of the speakers registered a pitch increase on the predicate. The most common dual pattern involved a pitch decrease on both the predicate and object (7/16). The predicate focus utterances were quite similar, with 1/16 speakers producing a pitch increase on the predicate, and 9/14 registering a pitch decrease on the object. Of note is that 7/16 Korean speakers registered a level change

between the focus operator and the predicate. When the two changes were analyzed together, there was no distinct pattern evident.

The speech patterns produced by the Spanish L1 participants were much closer to those of the native English speakers than the Korean speakers. In the object focus utterances, 15/16 recordings registered a pitch decrease on the predicate and 7/16 recording yielded a pitch increase on the object. However, in looking at the pitch changes combined, the results were bimodal for the most common pattern with 7/16 utterances showing a pitch decrease on both the predicate and object, and 7/16 utterances showing a pitch decrease on the predicate with a subsequent increase on the object. However, the contours for the predicate focus utterances didn't vary much from the object focus utterances. Only 2/16 utterances showed a pitch increase on the predicate and 10/16 utterances showed a pitch decrease on the object. The most common overall pattern involved a double pitch decrease on both the predicate and object (7/16).

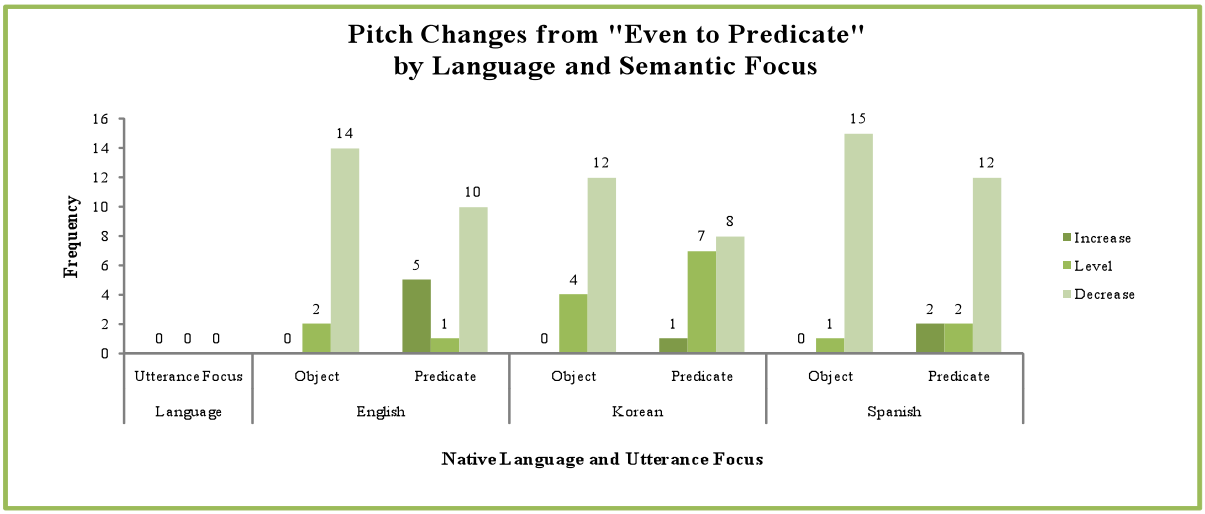

Figure VIII: EVEN - Pitch Changes from "Even to Predicate" by Language and Semantic Focus

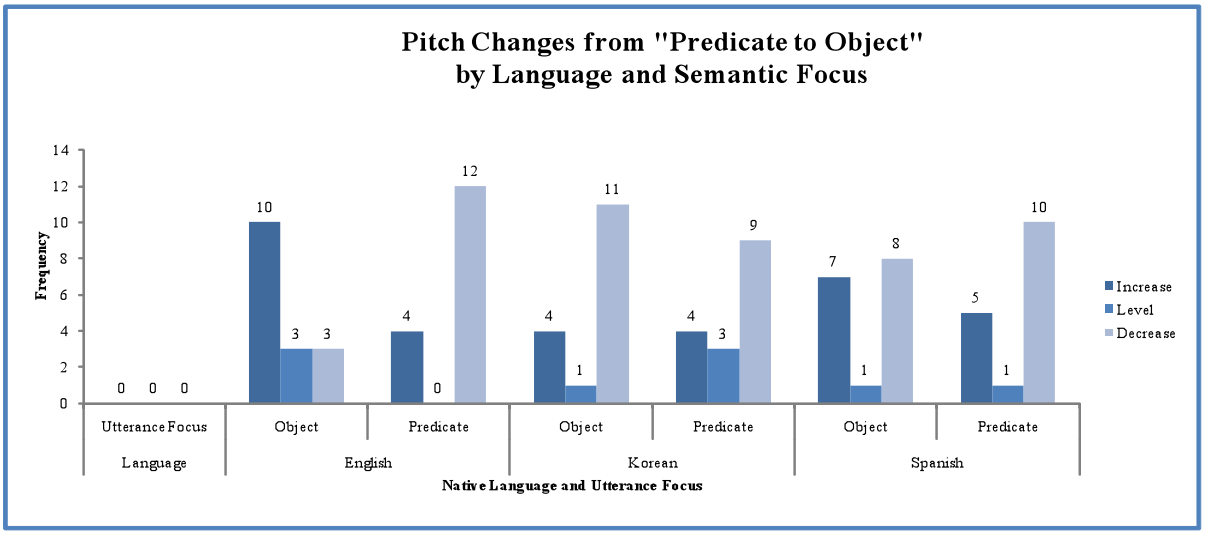

Figure IX: EVEN - Pitch Changes from "Predicate to Object" by Language and Semantic Focus

As with the focus operator "only", I conducted a series of Fischer Exact Probability Tests using 2x2 Contingency Tables to determine if the differences in pitch rises and falls were significant between the groups (See Appendix Table 6). Interestingly enough, the ratio of pitch increases to decreases from the focus operator "even" to the predicate was statistically identical (p=1) among the native English, Korean, and Spanish speakers in the object-focus contexts. The intonational similarities in predicate-focus contexts were also quite notable. However, there were statistically significant differences in pitch changes from "even" to predicate in object-focus contexts between Korean and native-English speakers (p=0.0213); this measure was almost significant between Spanish and native-English speakers (p=0.137). Of particular note, however, is that within native English speakers, there were significant differences in pitch change patterns between the object-focus and predicate-focus interpretations for both the "even to predicate" (p=0.042) and "predicate to object" measures (p=0.0009). This is important because it clearly indicates that native English speakers are using different prosodic patterns to resolve semantic ambiguities with the focus operator "even."

V.A.iii. “not-because”

For the scope of negation utterances involving the word “because”, 144 utterances were collected. An equal number was collected for the wide and narrow scope of negation contexts. To understand the different semantic interpretations of the “not-because” ambiguity, review example 7, replicated below. I hypothesized that speakers would produce some form of an increased pitch accent near the end of the phrase in the wide scope of negation contexts, and a decrease in pitch near the end of the phrase of the narrow scope of negation contexts.

(7a) BECAUSE>NOT (narrow scope of negation)

“Joseph usually loves to go out drinking. Tonight, however, **Joseph isn’t drinking because he’s unhappy.** We should talk to him.”

(7b) NOT>BECAUSE (wide scope of negation)

“Sometimes people drink when they’re depressed, but that isn’t the case tonight. **Joseph isn’t drinking because he’s unhappy.** He’s drinking because he just got a promotion.”

For each of the semantic scopes, 24 utterances were collected from each of the three native language groups (See in-text Figures X, XI; Appendix Tables 7, 8 for reference). Although I had originally decided to code the pitch level change from word of negation to the predicate, I chose to exclude this measure from my analysis. My reasoning for this exclusion is that a vast majority of the utterances (120/144) across contexts and native language groups registered a pitch decrease between these two points, and a mere 9/144 measured a pitch increase. The homogeneity of this result suggests that it would not be a useful measure for establishing a prosodic difference between these two contexts.

In the narrow scope of negation contexts of the native English speaker recordings, 6/24 utterances realized a pitch increase from the predicate to “because” and 19/24 utterances realized a pitch decrease from the verb to the focus of the dependent clause (the “cause”). Half of the utterances showed a decreasing pitch pattern from the predicate to both “because” and “cause.”

This finding fits the expected pattern regarding a decrease in pitch near the end of the utterance in the narrow scope of negation contexts.

In the wide scope of negation contexts of native English speakers, 12/24 recordings showed a pitch fall on “because”, while 9/24 recordings showed a pitch rise on the “cause.” The highest frequency pitch pattern also involved a pitch decrease on both “because” and “cause” (10/24). I had expected a general increase in pitch on the “cause” of the wide scope of negation utterances. While 9/24 does not reflect the majority of speech samples, there were notably more pitch rises on the “cause” in the wide scope contexts than in the narrow scope contexts. The general intonational differences between the two contexts, however, are not as evident as I had anticipated.

In the narrow scope utterances of the native Korean speakers, a mere 3/24 recordings yielded a pitch increase from the predicate to the “because”, but 18/24 recordings showed a pitch decrease on the focus of the dependent clause. Half of these recordings showed an overall pattern of a pitch decrease on both the “because” and the “cause.” In the wide scope contexts, the pattern of pitch contour shifted, but in the direction opposite that of the changes in the native English speaker recordings. There were no pitch increases on the “cause”, and 6/24 recordings yielded a pitch increase on “because.” This finding clearly contradicts the expectation of an increased pitch level on the “cause” in the wide scope of negation contexts. In contrast to the native Korean speakers, the L1 English speakers registered many more pitch increases on the latter end of the utterance in the wide scope interpretations. For the Korean speakers in the wide scope interpretations, the most common overall pattern was also a continuous pitch decrease from the predicate to both “because” and the “cause.”

The narrow scope recordings of the native Spanish speakers followed a pattern almost identical to that of the Korean speakers. Again, a mere 3/24 recordings showed a pitch increase on “because,” while 19/24 recordings showed a pitch decrease on the “cause.” The most common overall pattern involved a double decrease from the predicate to both “because” and the “cause” (11/24). In the wide scope utterances, 6/24 utterances yielded a pitch increase on “because”, and 3/24 yielded a pitch increase on the “cause.” Interestingly, 10/24 of these utterances showed a level pitch change from the predicate to “because.”

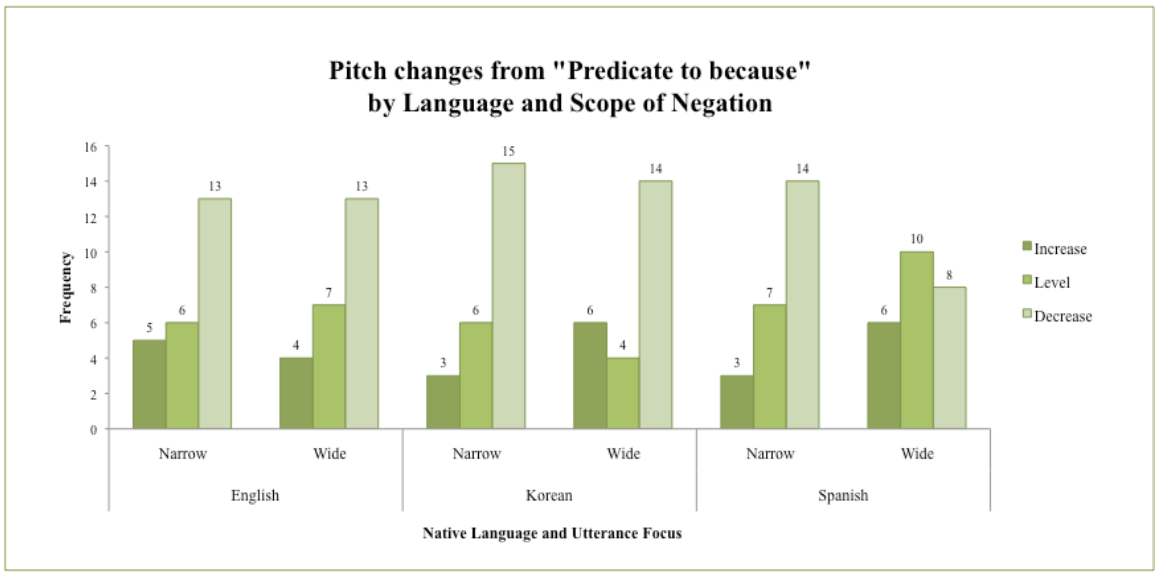

Figure X: NOT-BECAUSE - Pitch Changes from “Predicate to because” by Language and Semantic Focus

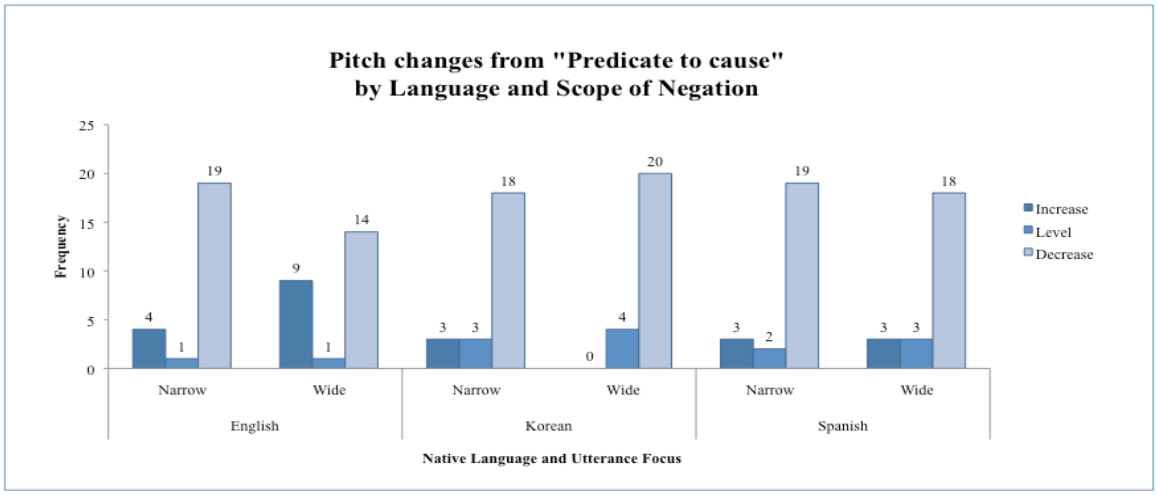

Figure XI: NOT-BECAUSE - Pitch Changes from “Predicate to cause” by Language and Semantic Focus

As with the focus operator utterances, I proceeded to perform a series of Fischer Projection Tests using 2X2 Contingency Tables to determine if the difference in the frequencies of pitch increases and decreases among the native language groups was significant (See Appendix Table 9). As a generalization, the differences between language groups and between the two scopes of negation were much less statistically different. For example, within native English speakers, the frequency of pitch changes from the predicate to "because" was statistically identical between the wide and narrow scope of negation. However, the difference in frequency of pitch changes from the predicate to the "cause" yielded a p-value of 0.189. While not statistically significant, a comparison of these two values suggests that further research should focus on pitch level changes occurring at the latter end of "not-because" utterances. Furthermore, the difference between Korean and English L1s from the predicate to "cause" in the wide scope interpretations was highly significant at p=0.002.

V.A.iv. "every"

For the utterances involving quantifier scope ambiguity, 144 recordings were collected. These were evenly split between utterances in which the quantifier held scope over negation and negation held scope over the quantifier. To illustrate these differences, example 9 is replicated below. I did not have a concrete hypothesis regarding the specific pitch differences that would be realized in these two scopes.

(9a) EVERY>NOT (quantifier scopes over negation)

> "Three lizards were riding around on motorcycles. They were trying to jump over a ramp. None of them made it. **Every lizard didn't jump over a ramp**."

(9b) NOT>EVERY (negation scopes of quantifier)

> "Three lizard were riding around on motorcycles. They were trying to jump over a ramp. The green and orange lizards were able to jump over the ramp. However, the black lizard was unable to jump over the ramp. **Every lizard didn't jump over a ramp."**

For each of these two scopes, 24 samples were collected for each native language group (See in-text Figure XII; Appendix Table 10 for reference). I originally collected pitch change data between the word containing the quantifier scope ("every", "everyone", and "everybody") and the word indicating negation ("not", "isn't", "didn't"). However, an overwhelming majority of the utterances across both scopes registered a pitch decrease between these two points (136/144). I have chosen to exclude this data from further analysis as it will not provide reasonable data to help tease apart the prosodic differences between the two contexts.

For the native English-speaker recordings in which negation has scope over the quantifier, 9/24 utterances registered a pitch increase from the word indicating negation (the "not) to the most prominent, stressed vowel of the remainder of the argument (the "argument"). Another 11/24 recordings registered a pitch decrease on the argument. In the utterances with quantifier scope over negation, there were also 9/24 utterances that registered a pitch increase on the "argument," and 12/24 utterances registered a pitch decrease. Statistically, these results are identical, which suggests that English speakers are not systematically disambiguating these contexts through prosodic means alone.

For the Korean speakers, 4/24 recordings yielded a pitch increase on the "argument" in both scopes, with 16/24 recordings showing a pitch decrease on the "argument" in the negation-scope utterances and 14/24 recordings showing such a decrease in the quantifier-scope utterances. Again, as with the native English speakers, the frequency of these results between the two scope contexts is statistically identical.

For the Spanish speakers, 5/24 recordings showed a pitch increase from the “not” to the “argument” in the negation-scope utterances, with 13/24 utterances showing a pitch decrease. In the quantifier-scope context, 3/24 recordings yielded a pitch increase on the “argument”, while 17/24 recordings registered a pitch decrease. The difference between the frequencies of these pitch changes is not as negligible as within the Korean and English speakers, but the two-tailed Fischer Exact p-value is 0.438, which is still far from significant.

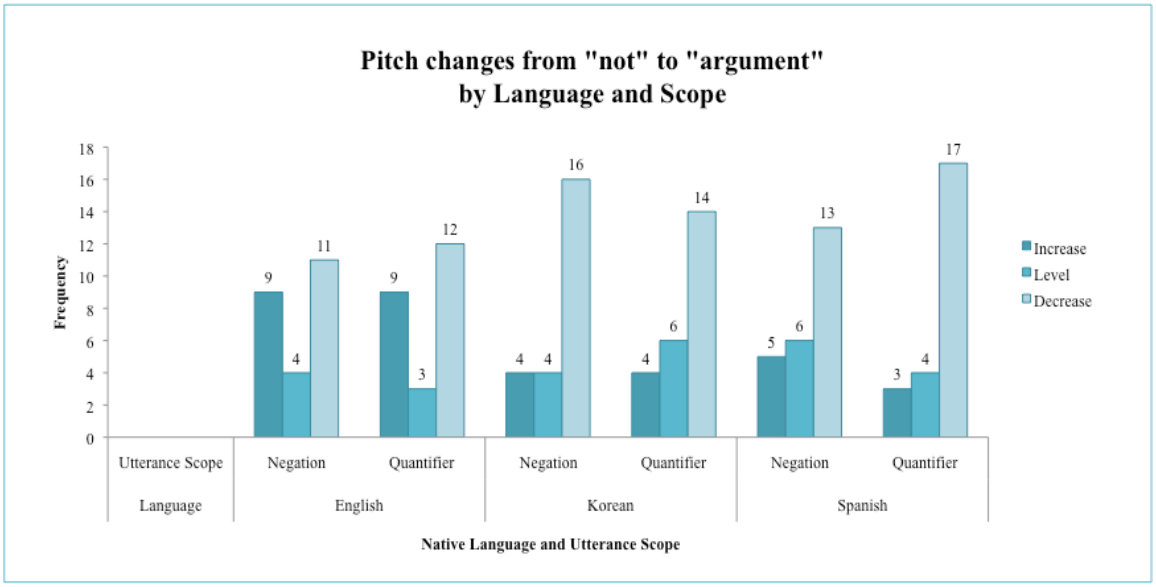

Figure XI: EVERY - Pitch Changes from “not” to “argument” by Language and Semantic Scope

Although there were not significant statistical differences between the two scope interpretations within speaker groups, the patterns of pitch change did differ somewhat between the L1 groups (See Table 11). In the negation-scope contexts, the difference in pitch frequencies between the English and Korean speakers approached significance at p=0.176. In the quantifier-scope contexts, the difference in pitch frequencies between the English and Spanish speakers also approached significance at p=0.085. While these differences aren’t significant within

groups, the inter-group differences suggest that the speakers are relying on different prosodic strategies for the disambiguation of quantifier scope.

V.B. Description of Perception Analysis

After finishing the perception study sessions, I compiled the participants' answers into SPSS software and ran a series of T-tests and ANOVAS to determine significance within and between groups. A participant's answer was determined to be "correct" if their perceptual choice aligned with the semantic focus of the target. First, I ran a number of One-Sample T-tests to analyze whether the different L1 groups had selected the correct target utterances at a proportion statistically greater than chance. In addition to analyzing their overall proportions correct, I also examined their results by the four types of semantic ambiguity (only, even, not-because, every), by the selection task (paragraph, speech), and by the semantic focus (object/predicate, wide/narrow). I also examined combinations of variables, which will be further described throughout the results. I later performed a series of One-Way ANOVAs to compare the means between L1 groups within these various categories. I then examined a series of paired sample T-Tests and Paired Sample Correlations, both overall and within L1 language groups, to determine if there were significant correlations between these various categories. At various points during the analysis, I combined the Korean and Spanish L1 speakers into an "ESL" category for the purpose of exploring overarching second language acquisition patterns.

V.B.i. "One-Sample T-tests and Averages"

The overall mean proportion of targets identified correctly among all speakers combined was 0.628, which, when compared to chance, was highly significant at $p<0.001$. When broken into categories, however, the results are much more complex (See Figure XIII for reference). The

mean proportions correct across all speakers by the four types of ambiguity were: “only”- 0.710 ($p<0.001$), “even”-0.530 ($p=0.600$), “not-because”-0.753 ($p<0.001$), and “every”-0.513 ($p=0.816$). This is critical because it shows that the participants performed significantly above chance for the “only” and “not-because” contexts, but essentially at chance for the “even” and “every” contexts. Their overall performance was statistically above chance for both paragraph and speech tasks, with a mean paragraph selection task of 0.616 ($p<0.01$) and a mean speech selection task of 0.640 ($p<0.001$). Their overall performance by focus and scope was also significantly greater than chance, with a mean object/narrow focus proportion of 0.576 ($p<0.05$) and a mean predicate/wide focus of 0.680 ($p<0.001$).

Among just native-English speakers, the significance trends were comparable to the sample as a whole, but these participants tended to perform better on an absolute basis within each individual category. Their mean overall proportion of correct target identifications was 0.722 ($p<0.001$), with significant performance in the “only” (0.861, $p<0.001$) and “not-because” (0.833, $p<0.01$) categories, but not the “even” (0.667, $p=0.081$) and “every” (0.556, $p=0.620$) categories. The native-English speakers also performed significantly greater than chance in both types of tasks, with an average score on the “paragraph selection” task of 0.711 ($p<0.01$) and an average score on the “speech selection” task of 0.733 ($p<0.001$). Furthermore, these participants performed above chance on both types of scope and focus ambiguity, averaging 0.656 ($p<0.01$) correct on the “object/narrow” contexts and 0.789 ($p<0.001$) on the “predicate/wide” focus utterances.

The native Korean-speakers performed drastically different than the native English-speakers in the perception study. Their mean overall correct target percentage was 0.569, and their results were not significantly greater than chance for any of the four types of ambiguity, for

either of the two tasks, and for either of the two scopes/focuses. Their highest mean target rates were in the “not-because” contexts (0.645) and the “only” contexts (0.625), which do follow the general patterns of relative performance among the various categories.

As a general statement, the native Spanish-speakers correctly identified the target utterances at a level between that of the native English and Korean speakers. Their mean overall proportion of correct target identifications was 0.581, which was not significant. However, their average scores for the “not-because” (0.770, $p<0.01$) and “only” (0.625, $p<0.05$) contexts were highly significant. Interestingly enough, these participants performed at a rate below chance for the other two types of ambiguity. These speakers did not select the correct targets at a level significantly greater than chance for either the “paragraph” or “speech” tasks. Of other particular note is that the Spanish speakers were able to correctly identify the “wide/predicate” focus utterances at a rate significantly greater than chance (0.650, $p<0.05$), but not the “narrow/object” focus utterances.

When the results of the Korean and Spanish speakers are conflated, a few slightly different patterns emerge. In addition to their combined mean target percentage being statistically significant (0.575, $p<0.05$), their results from the “only”, “not-because”, “speech task”, and “wide/predicate focus” contexts are all also statistically greater than chance.

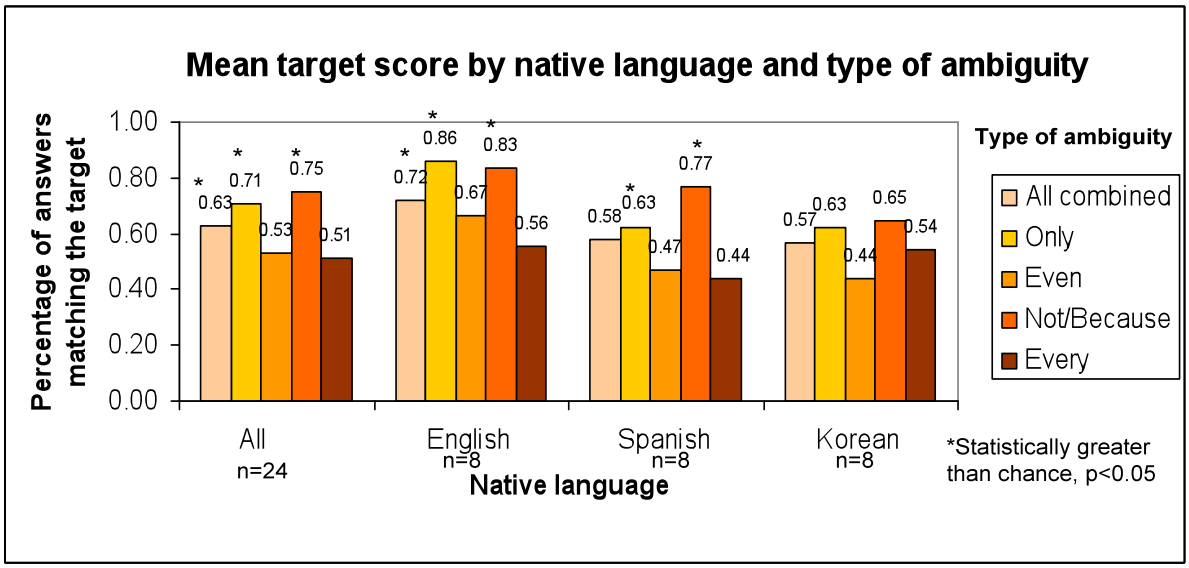

Figure XIII: Mean target scores by native language and type of ambiguity

V.B.ii. "ANOVA"

This next set of analyses examines the composite degree of difference found between the three language groups in their perceptions task results. A significant figure in this section indicates that there was a statistical difference when comparing performance within a given parameter among these three groups (See in-text Table 12 for reference).

The first set of ANOVAS refers to differential analysis between the native English-, Korean, and Spanish-speakers. The overall rates of correctly identifying the target sequences were significantly different among the three groups ($p<0.014$). Their differences in performance in the "only" category was also statistically significant ($p<0.021$); however their differences in performance were not significant in the other three categories of ambiguity. Their inter-group differences were also not significant in the two types of perception tasks. Furthermore, their differences were not significant in the perception of "object/narrow" utterances, but they were significant in the perception of "predicate/wide" utterances ($p<0.01$).

I then completed a set of ANOVAS in which participant answers were coded as either "native" or "non-native" English speakers. In other words, I compiled the results of the native Spanish- and Korean- speakers together and compared those results against the answers of the native English-speakers. The overall mean performance was statistically different between the two groups ($p<0.01$). As in the previous analysis, the difference was also different for the "only" ambiguity category ($p<0.05$), but not the other three types of ambiguity. However, there were statistical differences between these two groups for both the "paragraph" ($p<0.05$) and "speech" ($p<0.05$) selection tasks. There were also statistical differences between the performances of the two groups in the perception of "object/narrow" focus utterances ($p<0.05$) and "predicate/wide" focus utterances ($p<0.01$).

To satisfy my curiosity, I also performed a statistical comparison of means between the Spanish and Korean speakers. It is of note that their results are not significantly different from each other across any ambiguity category, selection task, or scope/focus interpretation. These results are particularly compelling because the native Spanish-speakers did perform statistically above chance for a number of categories in which the Korean-speakers failed to do so. However, this ANOVA suggests that the perception acuity of these two groups is not significantly different.

Table 12: One-Way ANOVA comparisons, significance values

	One-way ANOVA significance values		
	Comparing means across three groups	Comparing English L1 and L2 means	Comparing Spanish and Korean means
Overall	0.014**	0.003*	0.839
By Ambiguity Type			
only	0.021**	0.005*	1
even	0.192	0.068	0.83
not-because	0.107	0.107	0.164
every	0.668	0.583	0.448
By Task			
paragraph task	0.131	0.050**	0.593
speech task	0.069	0.021**	0.77
By Scope			
object/narrow scope	0.133	0.050**	0.667
predicate/wide scope	0.007**	0.003*	0.364
			*significant at p<0.01
			**significant at p<0.05

V.B.III. “Correlations”

I then conducted a series of paired sample t-tests to determine if performance results between individual categories were statistically different. Note that in these tests I was not comparing differences between speaker groups, but rather, between data categories. I jointly ran paired sample correlations between these multiple combinations of categories to determine the relationships between the various results (See Appendix Table 13).

With all participant data included, there were strong positive correlations between the overall target proportion and the performance within each of the four types of ambiguity. The highest paired-sample correlation was between the overall performance and “only” category,

with r=0.619 ($p<0.01$). This data suggests that participants achieving a high overall degree of target accuracy are also performing well within individual categories of ambiguity. However, when the perception results within each of these four categories were compared against each other, there were often significant differences, but not significant correlations. For example, the difference in participant performance between the "only" and "every" categories was significant at $p<0.01$, but the coefficient that correlated these two categories was a mere r=0.039 (p=0.854). There was neither a strong statistical difference nor a strong correlation between the two perception tasks. In contrast, there was both a significant difference ($p<0.01$) and strong significant correlation (r=0.566, $p<0.01$) between the results of the two focus/scope interpretations.

I then examined these categorical correlations within individual language groups, and unlike much of the other data, the patterns were quite different than what I found with the aggregate sample. Within just native-English speakers, the only significant difference in the ambiguity-category results appeared between the overall target rates and the "only" ambiguity category ($p<0.05$). The only significant correlation was realized between the overall target and the "every" ambiguity category (r=0.694, $p<0.05$); this statistic is particularly intriguing because native English speakers did not perform well in their attempts to perceive the targets of the "every" utterances. The overarching lack of significant correlations and differences suggests that an individual native-English participant was likely to perform more evenly across all four kinds of ambiguity. In contrast, there was both a strong statistical difference ($p<0.01$) and strong statistical correlation (r=0.849, $p<0.01$) between the native English speakers' ability to correctly perceive the "object/narrow" and "wide/predicate" focus interpretations.

Within the native Korean-speakers, there were virtually no statistically significant differences or correlations between the various utterance categories. However, considering that this group did not perform at a rate greater than chance within most of the categories, it is not surprising that there are no strong intra-group relationships between the various data groups.

The responses from the native Spanish-speakers, however, did reveal a number of curious quantitative relationships. Between the overall target performance and the "even" category there was both a significant difference ($p<0.05$) and a strong, positive paired samples correlation ($r=0.826$, $p<0.05$). There was also a significant difference between the overall target performance and the "not-because" category ($p<0.01$). Interestingly, there was both a significant difference ($p<0.01$) and a strong positive correlation ($r=0.724$, $p<0.05$) between the target percentages of the "only" and "not-because" categories. However, there were neither significant differences nor correlations between the two selection tasks nor between the two types of focus/scope.

V.C. Best Perception Answers

In addition to analyzing the perceptual results of individual participants and the statistical relationships between variables, I decided to examine the prosodic features of the individual questions on which participants performed best. When I initially selected the specific recordings to be used in the perception study, I had not yet conducted an analysis of the production results. Therefore, for the two interpretations of each of the twenty semantically ambiguous utterances, I selected the speech sample that I personally believed to be "the most evident" of the contexts. In taking a second look at these speech recordings, there are surprisingly few prosodic patterns among the samples that I selected. However, the inclusion of these particular samples in the

perception study is still valid because they were taken directly from the comparable semantic contexts in the production study.

There was one particular target which 100% of participants identified correctly. This question was a “not-because” speech-selection example in which speakers were given the wide scope of negation paragraph and both speech recordings. The presented context was as follows: “Gene didn’t cry because she was hurt. She cried because she was laughing so hard.” Of the available recordings, one sample was taken from an individual who had read the wide scope of negation in the production study (as above), and the other sample from the narrow scope of negation context: “Gene is normally a very expressive person, but her emotions tend to freeze up in difficult situations. Today, for example, when David got really angry at her, Gene didn’t cry because she was hurt.” The pitch contours of the selected speech samples are shown below, in Figures XIV and XV. The differences between the two recordings are quite distinct. In the narrow scope recording (Figure XIV), there is a small pitch increase between the “negation” and “predicate”, a small pitch decrease between the “predicate” and “because”, and a small decrease between the “predicate” and the “cause.” By listening to the recording, this particular speaker also seems to insert a pitch boundary after the predicate “cry”, but I was unable to see this pause on the spectrogram. In the wide scope recording (Figure XV), there is a large pitch decrease between the “negation” and the “predicate”, followed by subsequent smaller decreases in pitch between the “predicate” and “because”/“cause.” This speaker also places the highest pitch point on the subject “Gene”, which is quite different from the other samples. Again, using auditory cues, this speaker seems to utter the statement in one complete phrase without intermittent phrase boundaries, but it is difficult to visualize such information on the Praat spectrogram. When participants were presented with the wide scope of negation paragraph, 100% selected the correct

target recording, the one that corresponds to Figure XV. However, when the participants were presented with the narrow scope of negation context, only 62% selected the correct recordings, the one corresponding to Figure XIV.

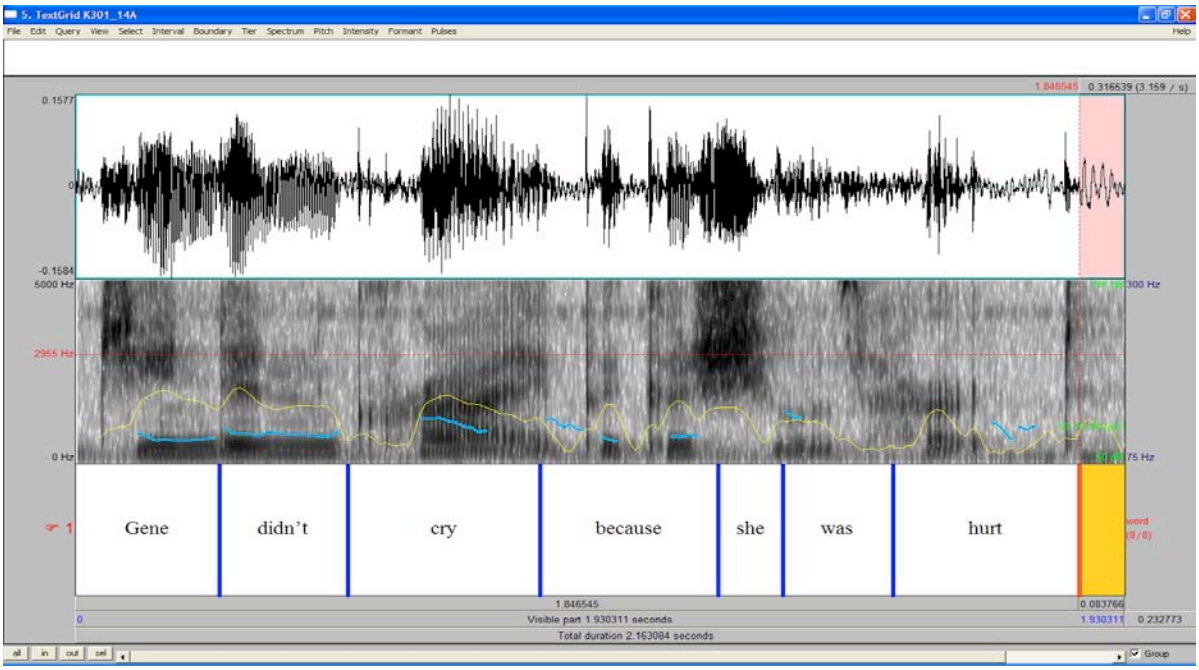

Figure XIV: Narrow scope of negation recording of "Gene didn't cry because she was hurt"

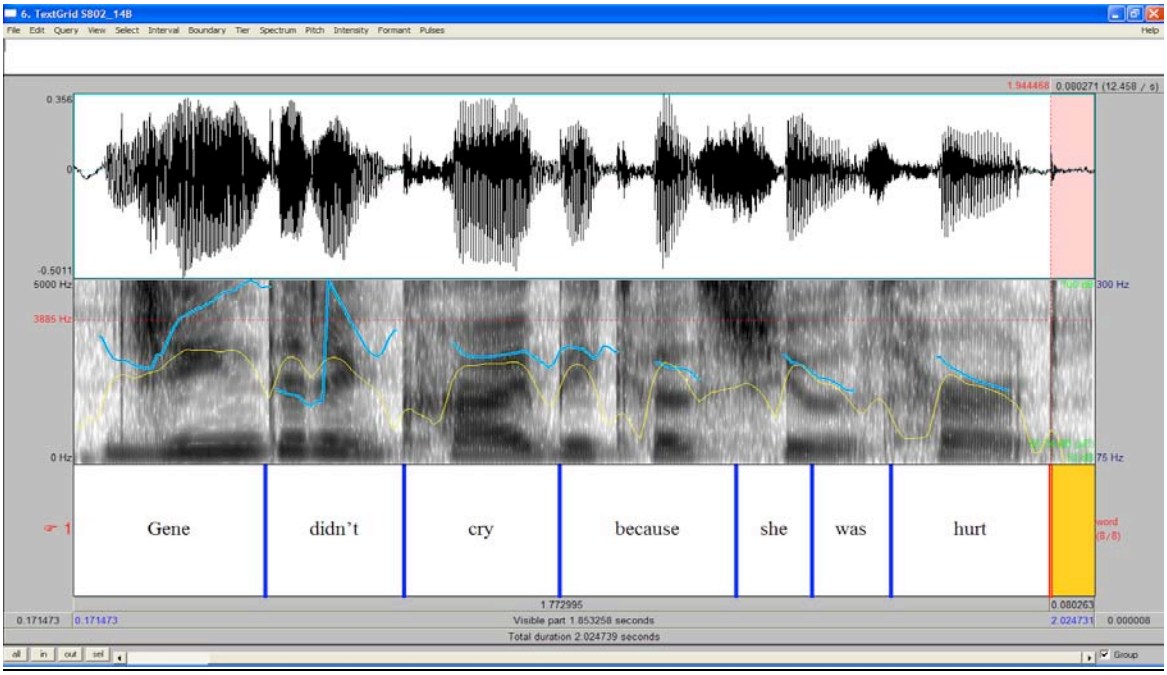

Figure XV: Wide scope of negation recording of "Gene didn't cry because she was hurt"

The overall perception results clearly indicate that participants discriminated among the prosodic cues available to them in selecting and differentiating between the targets. If they hadn't, then participants would have performed at a rate statistically equivalent to chance. However, upon reexamination of the selected speech recordings for the perception study, I am unable to find pitch patterns that predict the participants' likelihood of correctly identifying a given target. Essentially, the speech examples on which the participants "performed best" do not necessarily have the same pitch change patterns.

Of further interest is that there were certain recordings within the perception study that participants were remarkably likely to select, *regardless* of the correct target context. For example, one particular recording of a predicate focus "even" context was selected as the target speech sample 86% of the time, even though it was only the correct target 50% of the time. This is somewhat perplexing in that the alternate, narrow scope recording is quite discernibly distinct, both by auditory cues and the visual spectrogram. Paralleling this result to a lesser degree, within the "not-because" contexts, the wide scope of negation recordings were selected 61% of the time even though they were only the anticipated target 50% of the time. There must be some prosodic cue within these samples that inherently leads participants to select them. As far as the concept of prosodic disambiguation applies, however, we wouldn't expect listeners to select one speech sample as pertaining to both semantically ambiguous contexts. If semantic differences are encoded by prosodic means, and if such differences are perceived accordingly, then it would seem counterintuitive for one prosodic pattern to lead to both semantic interpretations. Luckily, these examples are exceptions to the norm.

IV. Discussion

The results from this study are dense. Instead of drowning in figures of statistical significance, I would prefer to highlight how these numbers relate to the original questions posed in this project. Namely, what role does prosody play in resolving semantic ambiguity within the English language in both listening and speaking, and if such a role exists, do speakers of other languages adopt this mechanism when learning English?

Prosody is an extremely complex linguistic concept. Based on the patterns of pitch change frequencies within the native English speakers, it is evident that intonation does supplement the process of semantic disambiguation; however, speakers do rely exclusively on prosody as mechanism for resolving this ambiguity. A key point to consider is that intonation has the potential to serve a host of different functions in language, such as indicating focus and stress, conveying implicature, and introducing new information, but it is not the sole linguistic device that can be used for accomplishing such speech acts (Hirschberg 2004:12). Furthermore, reliance on prosody for these linguistic functions varies across languages and between speakers within the same language. Nevertheless, based off the results from the production study, it is clear that for some ambiguous contexts, especially those involving focus operators, some native English speakers are employing varying pitch patterns to distinguish between semantic meanings. In the accompanying perception study, native English speakers were able to select the correct semantic target at a statistically significant rate without relying on contextual cues. Consequently, the importance of prosody as a key supplementary mechanism for clarifying ambiguity is undeniable.

In keeping a critical perspective, however, these studies also reveal that prosody is not an all-encompassing solution to this problem of inherent semantic ambiguity. Much of the data from

these two studies are “statistically greater than chance”, but the overarching patterns do not apply to every individual within the sample. From a theoretical stance, it makes sense that there would be multiple mechanisms for resolving a construct of ambiguity that is so engrained within the linguistic system. The results from the L2 English speakers indirectly lead me to infer that there must be linguistic tools in addition to prosody that are involved in the disambiguation process. Why? In the production study, participants in all three language groups were able to correctly answer virtually all of the forced choice questions that corresponded to the given contexts. Whether or not the speakers disambiguated those meanings prosodically, they nonetheless understood the correct meaning derived from the ambiguous statements. Furthermore, from the results of the perception study it is clear that speakers of all three languages understood that these semantically ambiguous statements had the capacity to yield multiple meanings in different contexts. However, the L2 English speakers did not distinguish between the meanings by prosodic means in either the production or perception study to the same degree as the native English speakers. Given that they were able to comprehend this concept of ambiguity, they must rely more heavily on alternative means to interpret meaning in semantically ambiguous contexts.

The differences in results between the native and non-native English speakers yield some interesting notions on the concept of language acquisition and prosodic transfer. It is of note in this discussion that the average length of time with English language experience was 10 years for the L2 English speakers; these individuals were working in a rigorous academic environment that generally assumes proficiency in English upon matriculation. Moreover, the L2 speakers clearly grasped the complexity of the ambiguity presented to them. However, their speech contours did not mirror the production patterns of the native English speakers; the speakers did not systematically disambiguate between semantic meanings through pitch variations, and they

did not interpret target meanings using prosodic cues with the same accuracy as native English speakers. Curiously enough, within the perception study, there were also no strong correlations between performance and number of years experience with English. This suggests that long-term English learners are developing a mastery of the English language without necessarily acquiring the same prosodic structure as native English speakers.

In discussions of cross-language similarities and differences, some linguists have posited that meaning conveyed by suprasegmental speech tools, such as intonation, is more likely to be shared across languages and should, therefore, be transferred more easily (Chen 2009: 368). The results of the current study, however, suggest that prosodic cues for the resolution of semantic ambiguity in English are neither transferred in production nor fully utilized in perception by L2 English speakers. Of course, the function of pitch accents is highly variable between languages. In order to develop a fuller understanding of the nature of this pattern of transfer, it will be necessary to examine how speakers resolve comparable semantically ambiguous utterances in their native languages. Due to the diverse nature of intonation across languages, recent research in second language acquisition has emphasized the need to include instruction on intonation patterning (Chen 2009: 367). The results of the current study affirm the importance of this notion in that the second language learners failed to acquire the prosodic patterning of the native English speakers in the observed semantic contexts. It is possible that since not all native English speakers utilize prosody as a (potentially redundant) tool for disambiguation, the intonation patterns are not being reinforced in the ears of the language learner. I expect that the degree to which any feature of a second language is acquired is dependent on consistency and degree of exposure to a stimulus. If not all native English consistently employ a prosodic mechanism for

disambiguation in these contexts, then it would reasonably follow that L2 English learners are not developing comparable pitch contour patterns.

Unfortunately, the explanation is not so simple. Why is it that participants in the perception study were most accurate in identifying the interpretations in the "not-because" contexts when their production patterns were so inconsistent? Prosody should, in theory, only be relevant to communication if is both encoded by the speaker and comprehended by the listener (Krahmer 2002:1). If listeners are accurately relying on some cue to interpret a particular meaning, then they should also be consistently encoding that cue in their speech to elicit the same meaning. As one researcher describes it, "prosodic cues depend upon speakers' knowledge of the situation: speakers provide prosodic cues when needed; listeners use these prosodic cues when present" (Snedeker 2003:103). It is possible that prosodic cues *are* being encoded by speakers in the "not-because" contexts, but that they have little to do with pitch differences; with more time, I would attempt to explore the production data in search of other elements of phrasing, such as pause insertions, boundary tones, and amplitude changes, that might provide the answer to this enigma.

I have a few other hypotheses for why participants performed inconsistently across the four types of semantic ambiguity. By their very essence, focus sensitive operators must take a focus. Since the English language often recruits pitch accents to indicate stress and focus, it might be more linguistically routine for English speakers to resolve focus ambiguity using pitch. In contrast, the "not-because" combination is a very specific linguistic context. When occurring on its own, negation does not necessarily yield ambiguity, nor does the word "because." Since these words occur with much higher frequency on their own than in combination, it makes sense that production of prosodic cues for disambiguation purposes is less consistent. The same

argument could be used for the ambiguity that arises between negation and quantifier scopes. "Every" and "not" occur more often on their own than paired within the same utterance. Alternatively, the correct interpretation in the "not-because" contexts is more crucial for complete comprehension than in focus operator contexts. As described more fully in the introduction, the denotation of the two potential meanings is essentially the same with focus operators, but the two potential meanings in "not-because" contexts are mutually exclusive. It is of greater urgency, therefore, for a speaker to use all available linguistic resources at their disposal to comprehend the correct meaning in the "not-because" contexts. In normal discourse, speakers will engage the surrounding context to mine the intended meaning; with limited stimulus, as in the perception study, listeners are obliged to rely more heavily on prosodic means.

I also have a potential suggestion for why participants neither encoded nor comprehended differences with quantifier-negation ambiguity. In English, speakers will naturally use the semantically ambiguous "every-didn't" phrases to convey both semantic meanings in their respective contexts. However, English has alternative phrase structures for each of the two possible meanings with quantifier ambiguity (see Example 17).

(17) Ambiguous utterance: **"**Every lizard didn't jump over the ramp."

(a) EVERY>NOT (quantifier scopes over negation)
Alternative gloss: "No lizard jumped over the ramp"

(b) NOT>EVERY (negation scopes of quantifier)
Alternative gloss: "Not every lizard jumped over the ramp."

Since speakers already have this context-dependent mechanism available, it is perhaps inefficient for prosody to supplement the disambiguation process. The existence of an alternative means to disambiguate these contexts does not truly explain away the dearth of prosodic differences found in this study. As previously mentioned, language trends often lean towards the redundant. Even if there is an alternative method for conveying a message, that in itself does not

excuse prosody's role (or lack thereof) in resolving the semantically ambiguous utterance. However, as language shifts and changes, if speakers become increasingly reliant on one particular phrase for communicating a message, such as the "alternative" phrasing in example 9, then it possible that the supplementary role of prosody in the original phrasing could disappear.

With the advent of speech recognition technologies, the concept of utilizing prosody to interpret semantic meaning is becoming increasingly relevant (Johan 1995:2). For example, VERMOBIL, a German telecommunications company, has invested resources to tease apart the specific prosodic tools associated with processing semantic ambiguity. The world is relying on English as a global lingua franca. In an era where instant communication is an expected element of ordinary human interaction, the study of semantic ambiguity and its associated linguistic counterparts within the English language are poignantly appropriate. The concept of linguistic ambiguity exists within all language, but the contexts in which it manifests itself may differ significantly between individual languages. The mere presence of this complex, logical sequence of ambiguity is baffling; however, humans have an incredibly potent capacity to problem-solve. With that in mind, I believe that this project successfully approached one of many linguistic resources that English speakers employ to confront the counterintuitive ambiguity that exists within their language.

V. Works Cited

Adams, Corinne, and R. R. Munro. "In Search of the Acoustic Correlates of Stress: Fundamental Frequency, Amplitude, and Duration in the Connected Utterance of Some Native and Non-Native Speakers of English." *Phonetica* 35.3 (1978): 125-56. Print.

Avesani, Cinzia, Julia Hirschberg, and Pilar Prieto. "The Intonational Disambiguation of Potentially Ambiguous Utterances in English, Italian, and Spanish." *Proceedings of the XIII International Congress of Phonetic Sciences*. (1995): 174-177. Print.

Baltazani, Mary. "Intonation and Pragmatic Interpretation of Negation in Greek." *Journal of Pragmatics* 38.10 (2006): 1658-76. Print.

Baltazani, Mary. "Quantifier Scope and the Role of Intonation in Greek." University of California Los Angeles, 2002. Print.

Beckman, Mary E., Julia Hirschberg, and Stefanie Shattuck-Hufnagel. "The Original Tobi System and the Evolution of the Tobi Framework." *Prosodic Typology: The Phonology of Intonation and Phrasing*. Ed. Jun, Sun-Ah. New York City: Oxford University Press, 2005. 10-54. Print.

Chen, Aoju. "Perception of Paralinguistic Intonational Meaning in a Second Language." *Language Learning* 59.2 (2009): 367-409. Print.

Delattre, Pierre. "Comparing the Prosodic Features of English, German, Spanish and French." *Iral-International Review of Applied Linguistics in Language Teaching* 1.3-4 (1963): 193-210. Print.

Eckman, Fred R. "Markedness and the Contrastive Analysis Hypothesis." *Language Learning* 27.2 (1977): 315-30. Print.

Goad, Heather, and Lydia White. "Ultimate Attainment in Interlanguage Grammars: A Prosodic Approach." *Second Language Research* 22.3 (2006): 243-68. Print.

Hirschberg, Julia. "Pragmatics and Intonation." *The Handbook of Pragmatics*. Eds. Horn, Laurence R. and Gregory L. Ward. Malden, MA: Blackwell Pub, 2004. Print.

Hirschberg, Julia and Cinzia Avesani. "Prosodic Disambiguation in English and Italian." *Intonation: Analysis, Modelling and Technology*. Ed. Botinis, Antonis. Dordrecht: Kluwer Academic Publishers, 2000. 87-95. Print.

Hirschberg, Julia and Cinzia Avesani. "The Role of Prosody in Disambiguating Potentially Ambiguous Utterances in English and Italian." *Proceedings of the ESCA workshop on intonation*. (1997): 189-192. Print.

Hoffman, Robert T. "Some Ambiguities in the Study of Ambiguity." *Resolving Semantic Ambiguity*. Ed. Gorfein, David S. New York: Springer-Verlag, 1989. 205-22. Print.

Johan, Bos, Batliner Anton, and Kompe Ralf. "On the Use of Prosody for Semantic Disambiguation in Verbmobil." 1995. Print.

Jun, Sun-Ah. "Korean Intonational Phonology and Prosodic Transcription." *The Original Tobi System and the Evolution of the Tobi Framework*. Ed. Jun, Sun-Ah. New York City: Oxford University Press, 2005. 201-29. Print.

Jun, Sun-Ah. *The Phonetics and Phonology of Korean Prosody : Intonational Phonology and Prosodic Structure*. Outstanding Dissertations in Linguistics. New York: Garland Pub., 1996. Print.

Kelm, Orlando. "An Acoustic Study on the Differences of Contrastive Emphasis between Native and Non-Native Spanish Speakers." *Hispania* 70.3 (1987): 627-33. Print.

Krahmer, Emiel, and Marc Swerts. "Perceiving Focus." *Topic and Focus* 82 (2007): 121-37. Print.

Nambu, Satoshi and Yong-cheol Lee. "Prosody and Semantics of the Focus Particles Always and Only in Korean: Theoretical Implications from a Perception Experiment." *Proceedings of Western Conference on Linguistics.* (2010). Print.

Pitrelli, John, Mary Mechman, and Julia Hirschberg. "Evaluation of Prosodic Transcription Labeling Reliability in the Tobi Framework." *In proceedings of International Conference on Spoken Language Processing* (1994): 123-26. Print.

Pytlyk, Carolyn. "Interlanguage Prosody: Native English Speakers' Production of Mandarin Yes-No Questions." *Proceedings of the Canadian Linguistic Association Conference*. Ed. Print.

Ramirez Verdugo, Dolores. "Prosodic Realization of Focus in He Discourse of Spanish Learners and English Native Speakers." *Estudios Ingleses de la Universidad Complutense* 14 (2006): 9-32. Print.

Rasier, Laurent & Philippe Hiligsmann. "Prosodic Transfer from L1 to L2. Theoretical and Methodological Issues." *Nouveaux cahiers de linguistique francaise* 28 (2007): 41-66. Print.

Rodd, Jennifer M., M. Gareth Gaskell, and William D. Marslen-Wilson. "Modelling the Effects of Semantic Ambiguity in Word Recognition." *Cognitive Science* 28.1 (2004): 89-104. Print.

Schafer, Amy J., et al. "Intonational Disambiguation in Sentence Production and Comprehension." *Journal of Psycholinguistic Research* 29.2 (2000): 169-82. Print.

Silverman, Kim, et al. "{Tobi}: A Standard for Labeling English Prosody." *Proceedings of the International Conference on Second Language Processing.* (1992): 867-70. Print.

Snedeker, Jesse, and John Trueswell. "Using Prosody to Avoid Ambiguity: Effects of Speaker Awareness and Referential Context." *Journal of Memory and Language* 48.1 (2003): 103-30. Print.

Stockmal, Verna, Dace Markus, and Dzintra Bond. "Measures of Native and Non-Native Rhythm in a Quantity Language." *Language & Speech* 48.1 (2005): 55-63. Print.

Ueyama, Motoko, and Sun-Ah Jun. "Focus Realization of Japanese English and Korean English Intonation." *Proceedings of the 7th Japanese Korean Linguistics Conference*. Ed. Print.

Wennerstrom, Ann "Intonational Meaning in English Discourse: A Study of Non-Native Speakers." *Applied Linguistics* 15.4 (1994): 399-420. Print.

VI. Appendices

VII.A. Appendix A: Additional Data Tables

only

Table 1: Separate Pitch Changes from "Only to Predicate" and "Predicate to Object" by Language and Semantic Focus

		From only to predicate				From predicate to object			
Language	**Utterance Focus**	**Increase**	**Level**	**Decrease**	**Total**	**Increase**	**Level**	**Decrease**	**Total**
English	Object	3	0	13	16	10	0	6	16
	Predicate	7	0	9	16	4	1	11	16
Korean	Object	5	2	9	16	6	1	9	16
	Predicate	3	3	10	16	5	3	8	16
Spanish	Object	1	4	11	16	6	0	10	16
	Predicate	3	4	9	16	1	3	12	16

Table 2: Whole utterance speech contours combining both "Only to Predicate" and "Predicate to Object" pitch changes by Language and Semantic Focus

Language	**Utterance Focus**	**Increase-Increase**	**Increase Level**	**Increase-Decrease**	**Level-Increase**	**Level - Level**	**Level-Decrease**	**Decrease-Increase**	**Decrease-Level**	**Decrease-Decrease**	**Total**
English	Object	1	0	2	0	0	0	9	0	4	16
	Predicate	0	0	7	0	0	0	4	1	4	16
Korean	Object	0	0	5	1	0	1	5	1	3	16
	Predicate	1	0	2	0	1	2	4	2	4	16
Spanish	Object	1	0	0	0	5	0	5	0	5	16
	Predicate	0	0	3	0	0	4	1	3	5	16

Table 3: Significance values for differences in pitch-change distribution

Comparisons		**Significance (p values)****	
Pitch Measure A	**Pitch Measure B**	**Object Focus**	**Predicate Focus**
English Only-P	Korean Only-P	0.417	0.433
English Only-P	Spanish Only-P	0.613	0.434
Korean Only-P	Spanish Only-P	0.17	1
English P-O	Korean P-O	0.289	0.689
English P-O	Spanish P-O	0.289	0.333
Korean P-O	Spanish P-O	1	0.16
English Only-P	English P-O	0.029*	0.458
Korean Only-P	Korean P-O	1	0.673
Spanish Only-P	Spanish P-O	0.184	0.321
Language	**Context Comparison**	**Only-Predicate**	**Predicate-Object**
English	Object vs Predicate Focus	0.252	0.0731
Korean	Object vs Predicate Focus	0.676	1
Spanish	Object vs Predicate Focus	0.59	0.093

*significant at p<0.05

**Fischer Exact 2-tailed

even

Table 4: Separate Pitch Changes from "Even to Predicate" and "Predicate to Object" by Language and Semantic Focus									
		From only to predicate				From predicate to object			
Language	Utterance Fo	Increase	Level	Decreas	Total	Increas	Level	Decrease	Total
English	Object	0	2	14	16	10	3	3	16
	Predicate	5	1	10	16	4	0	12	16
Korean	Object	0	4	12	16	4	1	11	16
	Predicate	1	7	8	16	4	3	9	16
Spanish	Object	0	1	15	16	7	1	8	16
	Predicate	2	2	12	16	5	1	10	16

Table 5: Whole utterance speech contours combining both "Even to Predicate" and "Predicate to Object" pitch changes by Language and Semantic Focus											
Language	Utterance Focus	Increase-Increase	Increase-Level	Increase-Decrease	Level-Increase	Level - Level	Level-Decrease	Decrease-Increase	Decrease-Level	Decrease-Decrease	Total
English	Object	0	0	0	1	1	0	9	2	3	16
	Predicate	0	0	5	0	0	1	4	0	6	16
Korean	Object	0	0	0	4	0	0	4	1	7	16
	Predicate	0	0	1	1	2	4	3	1	4	16
Spanish	Object	0	0	0	0	0	1	7	1	7	16
	Predicate	0	0	2	0	1	1	5	0	7	16

Table 6: Significance values for differences in pitch-change distribution				
Comparisons		Significance (p values)**		
Pitch Measure	Pitch Measure	Object Focus	Predicate Focus	
English Even-P	Korean Even-P	1	0.35	
English Even-P	Spanish Even-P	1	0.39	
Korean Even-P	Spanish Even-P	1	1	
English P-O	Korean P-O	0.0213*	1	
English P-O	Spanish P-O	0.137	0.704	*significant at p<0.05
Korean P-O	Spanish P-O	0.45	1	**Fischer Exact 2-tailed
English Even-P	English P-O	0.000034*	0.704	
Korean Even-P	Korean P-O	0.106	0.36	
Spanish Even-P	Spanish P-O	0.006*	0.39	
Language	Context Comparison	Even-Predicate	Predicate-Object	
English	Object vs Predicate Focus	0.042*	0.009*	
Korean	Object vs Predicate Focus	0.429	1	
Spanish	Object vs Predicate Focus	0.224	0.71	

not-because

Table 7: Separate Pitch Changes from Predicate to "because" and Predicate to "cause" by Language and Scope of Negation

		From predicate to "because"				From predicate to "cause"			
Language	**Utterance Focus**	**Increase**	**Level**	**Decrease**	**Total**	**Increase**	**Level**	**Decrease**	**Total**
English	Narrow	5	6	13	24	4	1	19	24
	Wide	4	7	13	24	9	1	14	24
Korean	Narrow	3	6	15	24	3	3	18	24
	Wide	6	4	14	24	0	4	20	24
Spanish	Narrow	3	7	14	24	3	2	19	24
	Wide	6	10	8	24	3	3	18	24

Table 8: Whole utterance speech contours combining both Predicate to "because" and Predicate to "cause" pitch changes by Language and Semantic Scope

Language	Utterance Scope	Increase-Increase	Increase-Level	Increase-Decrease	Level-Increase	Level-Level	Level-Decrease	Decrease-Increase	Decrease-Level	Decrease-Decrease
English	Narrow	0	1	4	3	0	3	1	0	12
	Wide	1	1	2	5	0	2	3	0	10
Korean	Narrow	0	1	2	2	0	4	1	2	12
	Wide	0	1	5	0	2	2	0	13	1
Spanish	Narrow	1	0	2	0	1	6	2	1	11
	Wide	0	2	4	2	1	7	1	0	7

Table 9: Significance values for differences in pitch-change distribution

Comparisons		**Significance (p values)****		
Pitch Measure	**Pitch Measure**	**Narrow Scope**	**Wide Scope**	
English P-b/c	Korean P-b/c	0.691	0.724	
English P-b/c	Spanish P-b/c	0.691	0.441	
Korean P-b/c	Spanish P-b/c	1	0.487	
English P-cause	Korean P-cause	0.724	.002*	
English P-cause	Spanish P-cause	0.441	0.094	*significant at p<0.05
Korean P-cause	Spanish P-cause	0.487	0.232	**Fischer Exact 2-tailed
English P-b/c	English P-cause	0.47	0.244	
Koreaj P-b/c	Korean P-cause	1	0.020*	
Spanish P-b/c	Spanish P-cause	1	0.112	
Language	**Scope Comparison**	**Predicate-b/c**	**Predicate-cause**	
English	Narrow vs Wide Scope	1	0.189	
Korean	Narrow vs Wide Scope	0.454	0.232	
Spanish	Narrow vs Wide Scope	0.232	1	

every

Table 10: Separate Pitch Changes from "not" to "argument" by Language and Scope

		From "not" to "argument"			
Languag	**Utterance Scope**	**Increase**	**Level**	**Decrease**	**Total**
English	Negation	9	4	11	24
	Quantifier	9	3	12	24
Korean	Negation	4	4	16	24
	Quantifier	4	6	14	24
Spanish	Negation	5	6	13	24
	Quantifier	3	4	17	24

Table 11: Significance values for differences in pitch-change distribution

Comparisons		**Significance (p values)****	
Pitch Measure	**Pitch Measure**	**Negation**	**Quantifier**
English not-arg	Korean not-arg	0.176	0.307
English not-arg	Spanish not-arg	0.328	0.085
Korean not-arg	Spanish not-arg	0.709	0.686
Language	**Scope Comparison**	**Not-arg**	
English	Negation vs Quantifier Scope	1	*significant at p<0.05
Korean	Negation vs Quantifier Scope	1	**Fischer Exact 2-tailed
Spanish	Negation vs Quantifier Scope	0.438	

Table 13: Correlation R-values and Paired Samples Significance Tests Between Data Variables

	All Participants		**Just English**		**Just Spanish**		**Just Korean**	
Variable Combinations	Paired Samples Correlation R-value	Paired Samples T-Test Significance	Paired Samples Correlation R-value	Paired Samples T-Test Significance	Paired Samples Correlation R-value	Paired Samples T-Test Significance	Paired Samples Correlation R-value	Paired Samples T-Test Significance
Overall vs Only	0.619*	.022**	-0.106	0.036**	0.664	0.291	0.638	0.482
Overall vs Even	0.525*	.054**	-0.192	0.568	0.826*	.038**	0.549	0.257
Overall vs Not-Because	0.504*	.001*	0.572	0.062	0.639	.003*	0.085	0.336
Overall vs Every	0.516*	.026**	0.694**	0.102	0.7	0.051	0.377	0.782
Only vs Even	0.368	.004*	-0.632	0.133	0.48	.049**	0.481	0.142
Only vs Not-Because	0.224	0.400	-0.004	0.724	0.724**	.006*	-0.171	0.877
Only vs Every	0.039	.009*	-0.038	0.032	0.102	0.080	-0.08	0.583
Even vs Not-Because	0.214	.001*	0.121	0.126	0.487	.003*	-0.115	0.198
Even vs Every	-0.203	0.849	-0.581	0.533	0.447	0.718	-0.33	0.592
Not-Because vs Every	-0.156	.003*	-0.124	0.066	-0.018	.013**	-0.334	0.483
Paragraph vs Speech	0.122	0.600	-0.119	0.738	0.156	0.882	-0.245	0.544
Object/Narrow vs Predicate/Wide	0.566**	.003*	0.849*	.000*	0.124	0.128	0.481	0.442
		Pearson Coefficient		Pearson Coefficient		Pearson Coefficient		Pearson Coefficient
							*significant at p<0.01	
							**significant at p<0.05	

VII.B. Appendix B: Experimental contexts

"only"

1A. Anna was angry at multiple people that day. She usually hits people when that happens, but this time she only hit Joseph.

1B. Joseph has been such a bad boyfriend to Anna. After he cheated on her, I'm surprised she only hit Joseph.

1Qa: Did Anna hit more than one person?
1Qb: Did Anna kick Joseph?

2A. Bill went to an awesome party last night. He talked to Sarah, Molly, Rob, and Cat. But it was weird. The next morning, he only talked to Sarah.

2B. Bill and Sarah went to the mall yesterday, and usually they have a whole lot of fun. Bill usually treats Sarah to dinner, takes her to the photo booth, and pays for the games at the arcade. But Bill didn't want to do much yesterday. He only talked to Sarah.

2Qa: This morning, did Bill talk to more than one person?
2Qb: Did Bill give Sarah a gift at the mall yesterday?

3A. Mrs. Cohen likes to host tea for her friends on Friday afternoons. She bakes up a storm. Usually the guests have home-made cookies, cakes, biscuits, and other baked goodies to choose from. Today, however, Mrs. Cohen only baked the cake. She ordered the other treats from a local bakery.

3B. Mrs. Cohen offers some pretty neat cake-baking workshops. She buys the cake ingredients, bakes the cake, spreads the icing, and then does all the fancy decorating. Today, however, her assistant did most of the work. Mrs. Cohen only baked the cake.

3Qa: Did Mrs. Cohen bake cookies and brownies?
3Qb: Did Mrs. Cohen decorate the cake today?

4A. Susannah is so stressed out about being busy all the time. I wish she would hold back on the caffeine. I tell her that the cafeteria offers a wide variety of juices, waters, and teas, but no. She only drinks the coffee.

4B. When Susannah lived in Colombia, she spent the summer on a coffee plantation. She helped plant the coffee beans, harvest the beans, grind the beans, package the coffee, and, of course, drink the coffee. Back in the US, the process is less involved. She only drinks the coffee.

4Qa: Does Susannah drink tea?
4Qb: Does Susannah grow coffee now?

“even”

5A. Jordan is a compulsive pyromaniac. He sets stuff on fire all the time. Yesterday, for example, he even burned a cat.

5B. Jordan has this weird obsession with hurting cats. He’ll often kick them and mess with their hair; today he even burned a cat. It kind of worries me.

5Qa: Has Jordan burned things before?
5Qb: Has Jordan hurt cats before?

6A. Rose was so hungry yesterday at the food festival. She’s usually a pretty picky eater, but she ate everything! She even ate bamboo.

6B. Rose had such a fun time at the specialty bamboo conference last spring. She built shelters with bamboo, chopped bamboo stalks, talked with other academics about bamboo stuff. Humans can’t actually digest it, but she even ate bamboo.

6Qa: Did Rose eat bamboo?
6Qb: Did Rose build stuff with bamboo?

7A. Jacob is so lucky! He works for a talent agency in Hollywood, and he communicates with movie stars all the time. He even talks to Julia Roberts.

7B. Jacob is a huge Julia Roberts fan. He watches all of her movies, buys all of her autographs online, and posts pictures of her throughout the apartment. When he dreams, he even talks to Julia Roberts.

7Qa: Does Jacob talk with many movie stars?
7Qb: Does Jacob talk to other movie stars?

8A. Julia has a super photographic memory. She memorizes absolutely everything. Before finals, she even memorizes our textbooks.

8B. Julia is obsessed with textbooks. She buys them brand new at the bookstore, copies them word for word, and then arranges them by the Dewey Decimal system that our librarian uses. She likes them so much that she even memorizes our textbooks.

8Qa: Does Julia memorize anything besides textbooks?
8Qb: Does Julia memorize comic books?

“not-because”

9A. Usually Joseph loves to go out drinking, but it was weird last night because he didn’t have a single beer. I just found out that Joseph isn’t drinking because he’s unhappy. We should talk to him.

9B. Sometimes people drink when they're depressed, but that isn't the case today. Joseph isn't drinking because he's unhappy. He's drinking because he just got a promotion.

9Q: Is Joseph unhappy?

10A. Maria eats chocolate when good things happen to her. She isn't eating chocolate now because she's depressed.

10B. Maria often overeats as a form of emotional therapy, but that isn't the case tonight. She isn't eating chocolate now because she's depressed. She's eating chocolate because her mom sent her a package of goodies.

10Q: Is Maria depressed?

11A. Laura is so sensitive to caffeine. It keeps her awake if she drinks it after 11 AM. But she drank some at 3 P.M. for some reason. She isn't sleeping because she drank coffee.

11B. Coffee has this really weird effect on Laura. It's supposed to keep people awake, but it makes Laura sleepy. Tonight though, she isn't sleeping because she drank coffee. She's sleeping because she pulled an all-nighter yesterday.

11Q: Did Laura drink coffee?

12A. Harold really dislikes his literature class, and after his four-hour swimming practice today, he just doesn't have the energy to deal with Herman Melville. Harold isn't reading because he's tired.

12B. Harold isn't reading because he's tired. Harold is reading right now because he has a midterm tomorrow.

12Q: Is Harold reading?

13A. Suzanne didn't run because she ate too much. She gets really bad cramps if she exercises after eating.

13B. Suzanne didn't run because she ate too much. She ran to stay healthy, just like she does every evening.

13Q: Did Suzanne eat too much?

14A. Gene is normally a very expressive person, but her emotions tend to freeze up in difficult situations. Today, for example, when David got really angry at her, Gene didn't cry because she was hurt.

14B. Gene didn't cry because she was hurt. She cried because she was laughing so hard.

14Q: Was Gene hurt?

"every"

15A. Three lizards were riding around on motorcycles. They were trying to jump over a ramp. The green and black lizards were able to jump over the ramp. The orange one was unable to jump over the ramp. Every lizard didn't jump over the ramp.

15B. Three lizards were riding around on motorcycles. They were trying to jump over the ramp. However, none of them made it. Every lizard didn't jump over the ramp.

15Q: Did any of the lizards jump over the ramp?

16A. I sent out the party invitations last week, but everybody hasn't replied yet. Mary and John got back to me already, but I'm still waiting to hear from George and Donna. I wish people would be more prompt.

16B. I thought I sent out the party invitations last week, but everybody hasn't replied yet. Maybe I didn't actually send the email. Nobody has called!

16Q: Did anybody respond?

17A. In the 1300s, the Black Plague swept through the world, destroying about a third of the world's population. However, everybody didn't get sick. Many people stayed completely healthy for some reason.

17B. When we were in Guatemala, the H1N1 virus broke out. I'm glad I didn't catch the flu. Actually, in our little tour group, everybody didn't get sick. We must have been immune or something.

17Q: Did somebody get sick from the tour group?

18A. My upperclassmen friends said that Dr. Jones' class is easy, but a lot of people struggled in her class. A couple people scored in the high 90s, but every student didn't do well.

18B. My upperclassmen friends said that Dr. Jones' class is incredibly hard. They told me that every student didn't do well. They weren't kidding about the difficulty. The highest grade she gave was a C-.

18Q: Did anyone do well in the class?

19A. Our grandparents definitely pick favorites among their grandchildren. I received a couple of presents for my birthday, and so did my brother, but not my cousins. Everyone didn't receive birthday gifts.

19B. Our mom told us that we had behaved really badly this year, so we weren't going to receive birthday presents. She wasn't kidding. Everyone didn't receive birthday gifts.

19Q. Did someone receive birthday presents?

20A. I am obsessed with chocolate, but I found out recently that everyone doesn't eat chocolate. My friend Nina, for example, is allergic to it, although her parents eat chocolate all the time.

20B. My friend Nina doesn't like chocolate, which I don't understand. In her family, in fact, everyone doesn't eat chocolate. I think that's crazy… my family eats chocolate all the time.

20Q. Does someone in Nina's family eat chocolate?