Speakers' production and listeners' perception of scopally ambiguous sentences

in a discourse context

Abstract

Researchers have long sought to determine the strength of the relation between prosody and the interpretation of scopally ambiguous sentences in English (e.g., All the men didn't go). While Jackendoff (1972) proposed a one-to-one mapping between sentence-final contour and the scope of negation (falling contour: narrow scope, fall-rise contour: wide scope), subsequent researchers (e.g., Kadmon & Roberts, 1986; Ward & Hirschberg, 1985) disentangled the link between prosody and scope, arguing that prosody is informative about information structure, which is in turn informative about scope. Such an account predicts variability in production, despite the existence of a correlation. To date, we lack systematic evidence to bear on this discussion. Here, we present a production study and two perception studies aimed at investigating whether speakers and hearers recruit auditory cues (including, but not limited to sentence-final contour) to disambiguate such sentences. We show that while there is considerable variability in production, there are prosodic and acoustic correlates to sentence interpretation. Moreover, hearers successfully recruit these cues to arrive at the correct interpretation. In light of these results, we argue that psycholinguistic studies (including language acquisition studies) investigating participants' ability to access multiple interpretations of scopally ambiguous sentences should carefully control for prosody.

1. Introduction

Sentences such as the following in (1) are notoriously ambigous, and have been widely discussed in the linguistics literature.

- (1) <u>All</u> the men did<u>n't</u> go.
- (2) a. $\forall x. man(x) \rightarrow \neg go$
 - b. $\neg \forall x. man(x) \rightarrow go$

Under one interpretation – the one captured by the formal logic in (2a) – none of the men went. (Roughly translated, *For all x, if x is a man, then x didn't go.*) In this representation the universal quantifier *all* (represented by \forall) takes scope over VP-level negation. The second interpretation – the one in (2b) – indicates that it is *not* the case that all of the men went. Under this interpretation, the possibility is left open that either only some (but not all) of the men went, or none went, as in (2a) – although the 'some' interpretation is the more salient of the two. In this case, negation takes wide scope over the universal quantifier.

Jackendoff (1972), building on earlier work by Bolinger (1965), proposed that speakers consistently use prosody to disambiguate sentences such as these (his (8.159)-(8.162) collapsed). Specifically, Jackendoff argued that the interpretation in (2a) (where negation has narrow scope under the quantifier) is indicated with a sentence-final falling contour, as in (3). By contrast, the interpretation in (2b) (where negation takes wide scope) is indicated with a rising or fall-rise contour, as in (4) (which Jackendoff referred to as Bolinger's 'B Accent', although this is not entirely accurate, given Bolinger's intonational system).

(3)
$$\forall > \neg$$
 (none)
ALL the men didn't go.
('A accent')

(4) $\neg > \forall$ (not all) ALL the men didn't go.

('B accent')

For Jackendoff, this difference was encoded in the representation of these sentences, and should therefore be minimally variable (if at all).

Prosody has been claimed to be intimately tied to scope and therefore to sentence interpretation in languages such as German (cf. Büring, 1997, 2003; Jacobs, 1984; Krifka, 1998; Martí, 2001; Sauerland & Bott, 2002). For example, the sentence in (5) uttered with the fall-rise prosody is claimed to only express the scopal relation in (6b), and not the one in (6a).

(5) Alle Politiker sind nicht korrupt.

'All politicians are not corrupt.'

/ALLE Politiker sind NICHT\ korrupt

(6) a. *all > negation

b. negation > all

A somewhat similar pattern also appears to hold in Greek (cf. Baltazani 2002, 2003). In English, there is by now considerable evidence indicating that speakers and hearers can use prosody for sentence disambiguation. The findings cover a range of lexical items and types of interpretation: structural ambiguity (Price *et al.*, 1991; Speer, Crowder, & Thomas, 1993); pronominal reference (Akmajian & Jackendoff, 1970; Hirschberg & Avesani, 1998, 2000; McMahon, Pierrehumbert, & Lidz, 2004); parenthetical v. integrated content (Price *et al.*, 1991); the scope of focus-sensitive operators *only* and *even* (Hirschberg & Avesani, 1997, 2000); and the interaction of a *because* clause with negation (Cooper & Paccia-Cooper, 1986; Hirschberg & Avesani, 1998, 2000; Koizumi, 2009). Thus, while we know that prosody *can* play a role in helping speakers favor and hearers arrive at a given sentence interpretation, it remains an open question how tight this relationship between prosody and sentence meaning is in English. Moreover, a gap still

exists when it comes to addressing Jackendoff's seminal claims for quantificational sentences with negation in English from the perspective of both production and perception.

Although small-scale studies have courted an answer this question (cf. Jackson, 2006; McMahon, Pierrehumbert, & Lidz, 2004), they have largely come up empty handed. This may be due in large part not only to their small sample size, but also to two other factors, which are at the heart of the current research. First, Jackendoff's emphasis (and Bolinger's) was on differences between the two sentences with respect to their sentence-final contour (i.e., a difference in rising v. falling contour). But as Jackson (2006)'s data suggest, and we demonstrate here, prosodic and acoustic differences (a) reside in other locations in the sentences, such as quantifier duration, but (b) are also highly variable within and across speakers. They therefore exist, but are not localized at the sentence end, and can be hard to pin down. Second, since Jackendoff, a number of researchers placed the phenomenon squarely in the purview of pragmatics, arguing that prosody highlights information in the discourse context, which in turn is informative about scopal relations (cf. Baltazani, 2002, 2003; Kadmon & Roberts, 1986; Ladd, 1980; Ward & Hirschberg, 1985, a.o.). Studies that do not systematically manipulate aspects of the discourse such as the presence and type of scalar alternatives (Hirschberg, 1986; Ward & Hirschberg, 1985), the role of negation in the presupposition or assertion, and the Question under Discussion (Roberts, 1996) may not be in a position to uncover the relevant data to demonstrate a connection between prosody and sentence disambiguation for sentences involving quantification and negation.

In this paper, we present a production study and two perception studies aimed at investigating whether both speakers and hearers recruit prosodic patterns to disambiguate quantificational sentences with negation. To our knowledge, this is the first systematic investigation of this phenomenon in English – and therefore the first to present both a quantificational analysis of correlates to sentence interpretation in speakers' production of the sentences, and robust connections between prosodic form and sentence interpretation in hearers' perception of these sentences. In §2, we begin with Jackendoff's original proposal about the role of intonation in the disambiguation of sentences with negation and quantification, then develop further predictions about the target sentences, taking into account pragmatic accounts developed by researchers in subsequent years. In §3, we review a range of experimental studies that bear on this question, and identify the remaining gap in the literature as precisely the one that we are addressing. In §4, we present an extensive production study aimed at identifying auditory cues to sentence interpretation. In §5-6, we present two complementary perception studies using stimuli from the production study, demonstrating that hearers can use such cues to retrieve the intended discourse-supported interpretation. Finally, in §7, we summarize our overall findings and situate our conclusions in the bigger picture.

2. Theoretical Background

Concerning the sentence in (7), Jackendoff's (1972) said that it has "a contrast in meaning...produced by a difference in the choice of pitch accent [read, intonation]" (pg. 352).

(7) <u>All</u> the men did<u>n't</u> go.

When the reading in (8) is intended, the choice of 'accent' associates negation with the presupposition, giving rise to the presupposition and assertion in (9). That is, it was expected that some quantity of men didn't go, and what is asserted is that that quantity is *all* of the men.

(8)
$$\forall > \neg$$
 (none)
ALL the men didn't go.
('A accent')

Presupposition: $\lambda Q [Q]$ of the men didn't go] is well-formed/under discussion

Assertion: all $\in \lambda Q [Q \text{ of the men didn't go}]$

When the reading in (10) is intended, the choice of accent associates negation with the focused quantifier (*all*), giving rise to the presupposition and assertion in (11). That is, there is some expectation that some men went, and what is asserted is that that number is not all. Since it is only possible to have a number less than all, the favored reading is that *not all* (or upper-bounded *some*) of the men went.

(10)
$$\neg > \forall$$
 (not all)
ALL the men didn't go.
('B accent')
(11) B accent (rise): negation is associated with the focus (i.e., the assertion), not
presupposition
Presupposition: $\lambda Q [Q \text{ of the men went}]$ is well-formed/under discussion
Assertion: all $\notin \lambda Q [Q \text{ of the men went}]$

Jackendoff did not restrict this account to sentences with subject *all* and negation. Indeed, he offered a similar approach for other negation sentences with *all* or *many* in object position (cf.

(12)), because (cf. the unfortunate example in (13)), and focused constituents elsewhere.

- (12) I didn't see ALL of the men. (his 8.181-8.182)
- (13) Max doesn't beat his wife because he LOVES her. (his 8.185-8.156)

For example, Jackendoff claimed that in both versions of the ambiguous sentence in (14) (his

(6.137)-(6.139) collapsed),

"Fred, the focus syllable, has a high pitch. After the focus syllable there is an abrupt drop to low pitch, which is maintained until almost the end. The ends, however, are different: the (a) reading has the falling coda of an A accent and the (b) reading has the rising coda of a B accent...As the focus is shifted, the same patterns appear in the intonation / contour to the right of the focus. In case the focus is in the final word, the pitch contour is compressed, but still recognizable" (pp. 259-260).

(14) a. \underline{FRED} doesn't write poetry in the garden.

'It is Fred who doesn't write poetry in the garden.'

b. <u>FRED</u> doesn't write poetry in the garden.

'It isn't Fred who writes poetry in the garden.'

Crucially, Jackendoff made the clear point that "the difference between them always appears at the end of the sentence" (pg. 260).¹

Liberman & Sag (1974) were, perhaps not surprisingly, among the first to attempt a revision of Jackendoff's account, arguing that encoding the distinction in the logic is a dead end. They proposed instead that with sentences such as the ones discussed above, where negation takes wide scope, the speaker is questioning the addressee's assumptions. The result is that there is a 'contradiction contour', and a distinct 'terminal rise'. Their key example, presented in (15) with the intonation in (16) (taken from Ladd (1980)'s summary of their account), may be thought of as a rejection to an addressee who has claimed that elephantiasis is incurable.

(15) Elephantiasis isn't incurable

¹ At this point, the reader may have noticed that the pitch accent pattern on *Fred* in this example (or on *all* in the examples above, for that matter) is not (necessarily) the same under both interpretations. We will return to this point – one that we find rather crucial –shortly.

...16 ...

While Ladd (1980) agreed with the general direction of Liberman & Sag's account, he pointed out that their 'contradiction contour' and the 'fall-rise' contour are not the same phenomenon, and that while both result in a rising sentence contour, the *beginning* of the sentence is noticeabily different.

To illustrate this point, Ladd highlighted the contrast between the two rising contours by embedding the same sentence in two different discourse contexts (his (8)-(9)).

(17) A: I just found out I'm going to die of elephantiasis.

B: Elephantiasis isn't incurable. [You won't die from it.]

(18) A: I'm doomed – the doctor just told me I either have elephantiasis or rabies.

B: Elephantiasis isn't incurable [...but rabies is.] ti Elephan asis isn't incurab¹⁰
(fall-rise tone)

Ladd also pointed out that while the contradiction contour cannot be embedded (as Liberman and Sag had observed), a fall-rise contour can be. Directly relevant to sentences with subject *all* and negation is Ladd's observation that sentences with a monosyllabic subject (such as (19), his (15)) obscure this difference, since there is only one syllable to host the accenting pattern.

(19) John's not in Boston.

When produced with a pitch accent on 'John' and a rising sentence-final contour, this sentence could have two interpretations, highlighted by Ladd's paraphrases and possible continuations in (20).

- (20) a. John's not in BOSTON what are you talking about; he's right in the next room watching the tube.
 - b. JOHN's not in Boston it was Henry's turn to go this time.

The implication, then, is that sentences such as *All the men didn't go* could result in a rising pattern for different reasons. These reasons will become more varied, and the possible discourse contexts more complex, as we proceed. Having teased apart the two rising patterns, Ladd focused on the fall-rise pattern in connection with Jackendoff's claims.

Ladd (1980) shared Jackendoff's observation that there are intonational differences correlated with interpretation, but saw the effect as pragmatic, rather than semantic. According to Ladd's account, the primary message of fall-rise is *focus within a given set*, and the relation between negation and focus arises only because '*all* can't be a subset so it must mean *not all*' (pg. 161). Specifically, the use of a fall-rise contour by a speaker indicates that the "variable of the focus presupposition" is a member of a contextually-relevant set. Ladd's example dialogue in (16), presented here as (21) illustrates this point: there is the presupposition that B fed something, and B's assertion combined with the use of the fall-rise contour (indicated with Ladd's `notation) indicates that s/he fed the cat, and cats are members of the set evokes by A.

- (21) A: Did you feed the animals?
 - B: I fed the `cat.

That the sets are contextually relevant and can be ad-hoc is illustrated by other examples, such as (22) (Ladd's (24)).

(22) A: What would you think of getting [-] a dog?

B: A^{*}stove, maybe.

As Ladd points out, dogs and stoves may share membership in a set of things like 'things we can

afford' or 'material possessions we can allow ourselves' – sets the speaker and hearer could perhaps agree upon. Ladd's reasoning behind the *all* example is reminiscent of Jackendoff's, but in a different vein: because fall-rise focus on *all* cannot place it as a member of a superset (because there is no greater quantity than *all*), then the interpretation must be *not all*. Thus, the conclusion Ladd arrives at is that the there are multiple types of rising contours, which result from different presuppositions, and that in the case of the fall-rise contour in particular, its purpose is to signal contextually-relevant set membership information between the speaker and hearer.

Ward & Hirschberg (1985) present what is easily the most extensive pragmatic account of fall-rise contour to date. Before presenting their own account, they proceed step by step through detailed reasons why the previous accounts were insufficient. Here, we present a subset of their key examples (using their \/ notation to indicate the syllable hosting the fall-rise accent pattern). Ladd's conception of "focus within a given set" cannot be correct, they argued, because of examples illustrating that B does not have to evoke membership *within* the set mention by A. For example, in (23) (their (19)), A is asking about the route that B took, traveling in Philadelphia. The dialogue is felicitous, they argue, if B does not know that Walnut ends at 34th. What is not at issue is membership in the set of streets.

- (23) A: Did you go straight up Walnut?
 - B: To Thirty-\fourth/.

Fall-rise can also accompany a superset, as in (24) (their (18)).

- (24) A: Are you sending me mail?
 - B: I'm sending peo/ple mail.

And fall-rise can be inappropriate even given membership in a set, as in (25) (their (17)).

- (25) A: Did they have a boy or a girl?
 - B: #They had a \boy/.

Moreover, fall-rise intonation need not accompany set membership: the dialogue in (21) above is still interpretable with falling intonation on 'cat'.

Having abandoned Ladd's 'focus within a set', among other possibilities, Ward & Hirschberg introduce their proposal. Put briefly, the purpose of the fall-rise contour is to convey "speaker uncertainty" about the appropriateness of an utterance in a given context, and specifically about a salient relationship between discourse entities. This may involve set membership, but not necessarily. What matters is that the speaker who employs the fall-rise contour perceives there to be some possible scale, and uses this contour, because s/he is uncertain about (a) whether it is appropriate to evoke a scale at all, (b) if some scale is appropriate, which scale should be chosen, or (c) given a scale, which value (or scalar alternative) should be chosen. This account is closely connected with Hirschberg (1986)'s dissertation work.

Ward & Hirschberg note that because theirs is a pragmatic account in which the contour is indicative of speaker knowledge of contextually-relevant scales and scalar alternatives, a fallrise contour should not be tied to any meaning (or scopal relation between quantification and negation) in particular, and should not force negation to take wide scope over the quantifier. Two predictions thus follow, which are relevant to the target sentences in question. First, we should be able to observe a fall-rise contour *without* negation taking wide scope over the quantifier (that is, with *all* taking wide scope over negation). Second, we should be able to observe negation taking wide scope over the quantifier *without* a fall-rise contour (i.e., with a *falling* contour). Indeed, Ward & Hirschberg provide examples for both such cases. For the first case, they present the context in (26) in which B responds to A's utterance using a fall-rise contour.

- (26) A: The foreman wants to know which meeting some of the men missed.
 - B: All/ the men didn't go to the last one.

Ward & Hirschberg argue that in this case, B's use of fall-rise is an indication of speaker uncertainty; B does not know whether A wants to know which meeting at least some of the men missed, or whether a quantifier scale should be evoked. For the second case, they point out that the embedded clause in (27) can be uttered with a falling contour while favoring the negation > *all* reading.

(27) George said that everyone had left for the game by five, but I know that all the men didn't go that early.

This rejection of a one-to-one correspondence between prosody and interpretation and an appeal to a pragmatic account is echoed in Kadmon & Roberts (1986), whose key ambiguous sentence is presented in (28), with the prosody-interpretation correlations in (29).

(28) He does<u>n't hate most of the songs</u>.

- (29) a. falling contour, response to 'negative' question*most* > negation
 - b. fall-rise contour, response to 'positive' question negation > most

We note in passing that is possible to find similar attested examples in a current search online.

(30) Although they did<u>n't win many</u> awards during their performing years, Led Zeppelin's lasting influence has garnered them several Grammy Hall of Fame Awards, as well as an

induction into the Rock & Roll Hall of Fame.²

(31) AMC's superb "Mad Men" did<u>n't</u> win <u>many</u> awards on the night, but the two they did counted...³

Kadmon and Roberts acknowledge that the prosody of a given sentence does indeed appear to favor one interpretation over another, but these different prosodic patterns also differ with respect to the contexts in which they occur, and the question under discussion that they address (cf. (9), (11)). In a nutshell, their claim is the following, "Prosody does not directly determine the relative scope of operators. Intonation and stress convey partial information about the structure of the discourse, and it is this structure which determines the relative scope. Since prosody does not give complete information about the structure of the discourse, it does not disambiguate the scope relations" (pg. 18). That is, prosody may favor an interpretation out of the blue, but because it is only partially informative about the discourse structure, the hearer is left to reconstruct the best fitting, and simplest interpretation. The hearer does this in large part by retrieving the relevant (implicit) question, which is part of the information structure, from the preceding utterance context. This question is connected to a presupposition skeleton along the lines of Jackendoff (1972) in that the question is either 'negative' or 'positive' and the presupposition either encodes negation or not (in which case negation is focused).

What follows from this discussion is that while it is certainly *possible* that a fall-rise (or more generally, non-falling) contour may favor a reading where negation takes wide scope over

²'Led Zeppelin and the Recording Connection Audio School'

⁽http://www.recordingconnection.com/artists/led-zeppelin)

³ Sepinwall on TV: Recapping the Emmys'

⁽http://www.nj.com/entertainment/tv/index.ssf/2008/09/sepinwall_on_tv_recapping_the.html)

the subject quantifier, such a contour is neither necessary nor sufficient for this interpretation. A certain degree of variability should be expected, if the account of this phenomenon is pragmatic in nature, such that prosody is tied to speaker uncertainty in a discourse context and implicit questions that are part of the information structure. One wonders, then, how strong this correlation is – that is, how likely both speakers *and* hearers are to associate a falling contour with an interpretation where negation takes narrow scope and a fall-rise contour with an interpretation takes wide scope. These are the precisely the questions that motivated the current research.

3. Experimental Background

3.1. The disambiguating potential of prosody

A number of studies over the years have shed light on prosody's role in sentence disambiguation. These studies have covered a range of lexical and syntactic patterns, and have used a variety of paradigms. The general picture that emerges is that both speakers and hearers assign a key role to prosody in determining the interpretation of a potentially ambiguous string, but this effect is variable among speakers, hearers, and contexts, and a conspicuous gap exists when it comes to systematically probing the sentences that originally interested Jackendoff and subsequent researchers.

Based on the experimental results from a set of three studies, Speer, Crowder, & Thomas (1993) argue that "prosody is maintained not only when it determines the syntactic form of an otherwise ambiguous sentence, but also when it contributes more subtly to sentence meaning, by determining the focus or the presuppositional structure for a sentence" (pg. 354). In one study, participants heard a prerecorded potentially ambiguous sentence, then selected from two choices the appropriate paraphrase of it. Among the stimuli were sentences with ambiguous pronominal

reference, phrase boundaries (e.g., *The dog may attack (#) Gwen.*), conjunction patterns among NPs (e.g., *Either Sam # or Susan and Lara #* v. *Either Sam or Susan # and Lara # ...*) and stress altering syntactic structure (e.g., *They are cooking APPLES v. They are COOKING apples*). Participants were indeed sensitive to the prosodic pattern when choosing their paraphrase – more so for syntactic ambiguities than for placement of focus.

Price *et al.* (1991) recorded professional radio announcers reading a range of ambiguous "segmentally identical but syntactically different" sentences that were embedded in contexts. Among these were integrated phrases v. parentheticals or appositional phrases (e.g., *Does Ian know (,) I wonder (,) when he will leave; The neighbors who usually read [the dailies /, the Daleys,] were amused*), attachment ambiguities (e.g., *Raoul murdered the man with a gun*), and verb particles v. prepositional heads. They then presented the excised sentences in isolation to naïve participants, who were given a choice between two contexts in which the sentence could have appeared and asked to select the best match. Price *et al.*, found considerable variation among the speakers, as well as variability in participants' success based on the speaker and the sentence types. However, overall, participants were successful at using prosodic cues for disambiguation.

3.2. Ambiguous sentences with negation and because

Other studies have focused more specifically on scopal ambiguity with negation. For example, Cooper & Paccia-Cooper (1986) presented participants with ambiguous target sentences such as (32), followed by a disambiguating follow-up sentence or embedded in a disambiguating paragraph.

(32) Dick did<u>n't</u> fly the kite <u>because</u> it was a beautiful day.

Participants were made aware of the ambiguity in one experiment, but not in another. Only when

they were made aware of the ambiguity and when producing the *because* > negation interpretation did participants significantly lengthen the vowel of the word preceding the *because* clause and increase the prosodic break (i.e., the pause) before *because*.

Hirschberg & Avesani (1997, 2000) presented English and Italian speakers (n=6 in English) with a range of ambiguous sentences, including *because-negation* (e.g., *William isn't drinking because he's unhappy*), attachment ambiguities, focus-sensitive operators (*only, even*), and the scope of *none* (e.g., *The election of none of these candidates would be a disaster*). Speakers were shown two paragraphs, each favoring a different interpretation, side by side, and asked to read them both aloud. Afterwards, they answered a comprehension question. Speakers produced no consistent results for *none* or for attachment ambiguities. They did, however, seem to produce differences for *because-negation*, as captured in Table **1**.

Table 1: Pattern of results for because-negation sentences in Hirschberg & Avesani (1997, 2000)

scopal relation	'internal phrase boundary'	falling contour
<i>because</i> > negation	12/18	15/18
negation > <i>because</i>	2/18	8/18

The data concerning the presence of a prosodic break (the first column of data) are consistent with Cooper & Paccia-Cooper (1986). The contour data can also be re-cast by saying that of the 13 sentences in which there was no falling contour, 10 were observed with negation took wide scope over *because* – as would be predicted. For *only/even* sentences, participants also placed the nuclear stress in the focused phrased in the vast majority of cases (*only*: 34/36 sentences; *even*: 30/36 sentences).

Further evidence for the role of prosody in disambiguating *because*-negation sentences in English comes from Koizumi (2009)'s self-paced silent reading studies. In one experiment, participants were shown sentences that favored one of two possible interpretations, as in (33), then answered a question about the sentence.

(33) a. Jane didn't purchase the white blouse because it had a stain. (*because* > negation)

b. Jane didn't purchase the white blouse because it suited her. (negation > *because*) Koizumi found that sentences favoring a *because* > negation scopal relation were read more quickly than negation > *because* sentences (a pattern consistent with Frazier & Clifton, 1996). However, this difference in reading time went away when the sentences were embedded in an *if* clause, which carries with it its own continuation rise. Thus, because both types of sentences were accompanied by a rising contour in this syntactic context, the intonational difference between the two sentences was neutralized. When in a subsequent experiment, Koizumi inserted a line break before the *because* clause (which had the effect of inserting a prosodic break), the preference for the *because* > negation reading again surfaced. These combined findings are consistent with those of Hirschberg & Avesani (1997, 2000): the sentence-contour and the presence/absense of a prosodic break before *because* play a significant role in the disambiguation of sentences with *because* and negation.

3.3. Ambiguous sentences with negation and quantification

More recently, researchers have been concerned with speakers' production and hearers' perception of ambiguous sentences involving quantification and negation. Baltazani (2002, 2003)'s complementary production and perception studies in Greek offer the first strong suggestion that speakers can use prosody to disambiguate quantificational sentences such as the target sentences of interest in this research. Baltazani focused on three types of items, captured in English in (34).

(34) a. He's <u>not</u> watching TV <u>because</u> he's bored (*because*-negation)
b. They did not eat many apples. (negation and object quantifier)

c. <u>Three</u> nurses helped <u>every</u> doctor. (subject and object quantifiers) Each item type was instantiated by a number of tokens, including a small range of quantifiers and varied SVO/SOV word order (since Greek allows for variable word order). In the production study, participants (n=5-8 for each sentence type) viewed a question that either did or did not contain negation, thereby allowing the positive/negative status of the QUD to be manipulated. They then saw the target sentence offered as an answer to the question, and were asked to read the question-answer pair aloud.

These productions were then presented to participants in three different perception studies. The number of participants per study ranged from approximately 30 to 90. Participants heard the excised target sentence in isolation, and were asked to choose a suitable answer, given a forced choice.

- (35) A: How many problems [did/didn't] the students solve?
 - B: The problems they solved are not many.

Baltazani found that her Greek speakers reliably produced ambiguous sentences with negation with two distinct contours related to the two interpretations, and that hearers were able to correctly identify their corresponding interpretations. While sentences involving subject-object quantifier interaction (like (34c) above) were produced in a manner similar to those like (34b), hearers were not able to unambiguously select the corresponding interpretation based on production alone. The combined production-perception experimental findings indicate that Greek speakers *do* reliably disambiguate sentences by using prosody, and that hearers' ability to retrieve the correct interpretation based on prosody alone depends on lexical items interacting and giving rise to the ambiguity.

Note that although these findings are extremely promising, they were found in Greek – a

language that may require such prosodic manipulation, or at least be more highly sensitive to it than English. Thus, it still remains an open question whether such a pattern could be observed in English for the quantificational sentences in question here. Recent studies aimed at addressing precisely this question have come up relatively empty handed. However, it may be that because these studies were done on a small scale and had a more narrow focus, and did not systematically manipulate the discourse context in which these sentences appeared, they were not in a position to uncover such a pattern. We turn now to these studies.

McMahon, Pierrehumbert, & Lidz (2004) designed a set of four short children's stories, each of which included a sentence ambiguous pronominal reference and and ambiguous sentence with the universal quantifier *every* interacting with negation. The story context and a follow-up continuation either favored negation taking narrow scope (n=2) or wide scope with respect to the quantifier (n=2), as in (36).

- (36) a. "Every bunny didn't jump over the fence, not a single one jumped over," saidHenry. "I guess you're still too small to play with my car."
 - b. "Every bunny didn't jump over the fence, only some did," said Henry. "I guess you're still too small to play with my car."

Eleven parents read these stories to their children while being recorded. Sentences were later coded for rise/fall contour. Contrary to predictions based on a strict relation between prosody and scopal relation, there was no discernible pattern with respect to sentence-final contour: there was a consistent 58% falling contour across all sentences.

The excised target sentences were then included in a perception study. Participants were explicitly told during a training session that prosody could be used to disambiguate sentences. During the test session, they were given two pages from the story – one that introduced the plot

and characters, and one with a forced choice of images representing what had immediately happened before the target sentence was uttered. They were given an unlimited amount of time to inspect these stimuli before listening to the target sentence and making their selection. While participants were successful with the pronominal reference items, they were largely at chance with the quantificational sentences, selecting the negation> *every* interpretation 59% of the time, regardless of the interpretation intended by the speaker. Moreover, when the results were recast in terms of prosodic patterns, it became apparent that participants were not recruiting prosodic patterns, but instead displayed a preference for the negation> *every* interpretation (as can be seen from the skew towards the top row in Table **2**).

Table 2: Distribution of percentage of perceived interpretation based on sentence-final contour for *every*-negation sentences in McMahon *et al.* (2004)

perceived interpretation	rising contour	falling contour
negation > every	61	58
<i>every</i> > negation	39	42

These results, then, could be taken as demonstrating that there is in fact *no* connection between prosody and sentence interpretation. Such a conclusion would, however, fly in the face of the intuitions a number of researchers have voiced over the years, which we reviewed above, and stand in contradiction with the previous experimental findings in English and other languages. It is possible, however, that certain elements of McMahon *et al.* (2004)'s experimental design minimized the role of prosody. For example, the pictures clearly disambiguated the sentences, leaving little room for prosody to play a central role in production. The continuations following the target sentence (provided in (36) above) may also have played a significant role. First, the follow-up phrasing may have been able to host the informative pitch accenting, minimizing the need for such prosodic information in the target sentence itself. Moreover, the construction of the dialogue may have favored a continuation rise in both cases. Rather than concluding from this study, then, that speakers *do not* offer the hearer such prosodic cues (and that when such cues are minimal or absent, hearers cannot retrieve them), we can conclude that speakers do not *always* do so, and when they do, such cues may not be localized in the sentence-final contour.

Jackson (2006) conducted a production study, in which he presented four speakers with nearly 200 sentences involving the scopal interaction of a small set of lexical items in different syntactic positions. These lexical items included negation, the universal quantifier *every*, and the indefinites *a*, *a few*). Each sentence was accompanied by two still images, illustrating the two interpretations (e.g., a group of circles, each hitting a square v. a group of circles, all hitting the same square for the sentence *Every circle hit a square*). Participants were asked to favor one of the two interpretations over the other in their production. Jackson's findings hinted at durational differences in the lexical items being correlated with the interpretation favored by the speaker. However, the conditional ranking among these durations was quite complex and the number of speakers extremely small, making generalization beyond his results rather difficult. Moreover, there was very little room for manipulation of information structure with such stimuli—a point think is important.

Given the previous experimental studies, we are left with the following picture. First, speakers and hearers *do* recruit prosodic information when disambiguation a range of sentence types in English, as well as in other languages. The ability to do so, however, is dependent upon the speaker and the type of ambiguity. Second, for the sparse evidence we have for English sentences with quantification and negation – precisely those sentences that have driven the debate over the years – it is not at all clear whether and when speakers use prosody to

disambiguate, and if so, whether the sentence-final contour is the locus of this informative prosody. If speakers *can* be found to use prosody to disambiguate sentences in favor of one possible interpretation, there is still an open question regarding whether or not hearers can recruit this information in the service of arriving at the correct interpretation. Finally, because the few studies that have investigated this phenomenon in English have not incorporated into their experiments certain key elements (e.g., clear negative/positive question under discussion, role of negation in presupposition/assertion relation, scalar alternatives), it remains unclear how much information structure in the discourse context plays a role in the use of prosody and in the observed variability in productions. The current research was designed to fill this gap.

4. Experiment 1: Production Study

The aim of this study was to elicit productions of ambiguous sentences in a discourse context with the purpose of identifying auditory (i.e., prosodic and acoustic) correlates of sentence interpretation.

4.1. Experimental Method

4.1.1. Participants

26 undergraduates participated. In all studies reported in this paper, the participants were undergraduates who received course credit in an introductory psychology or linguistics course in exchange for their participation. Data from six participants were excluded for reasons of nonnative status. In addition, one speaker's sound files were damaged and therefore excluded from analysis. In the end, data from 19 participants (14 F, 5 M) were analyzed.

4.1.2. Stimulus Design

Test items were scopally ambiguous sentences involving either universal quantification in subject position, or *many* or *most* in object position, and negation. There were multiple examples

of each test sentence type, presented in more than one discourse context, in which the information structure was varied. There were four contexts for the universal quantifier and three for *many* and *most*, for a total of 16 universal quantifier items and 12 *many* and *most* items. These 28 test items were then pseudorandomized with 28 control items, for a total of 56 experimental items. A full list of experimental items is included in Appendix A. In constructing the sentences, we controlled for sonorance, particularly at the end of the sentence and in the test items, in order to elicit as smooth a pitch track as possible.

Control items were also ambiguous sentences, which have been shown in previous research to be able to be disambiguated through prosody (see §3 above). These items included five pairs sentences with a *because* clause interaction with negation, five pairs of sentences containing a focus-sensitive operator (three with *only* and two with *even*), and four pairs containing ambiguous pronominal reference. Each individual control sentence was presented in two different contexts, each favoring one of the two competing interpretations. An example of each test and control item type is included here.

<u>All the magnolias won't bloom.</u> (universal quantifier, negation) (37) Liam doesn't know many alumni. (many, negation) (38)(39) Neil doesn't enjoy most musicals. (*most*, negation) (40)Georgia isn't singing because she's preparing for an audition. (*because* clause, negation) (41) Warren only likes the Orioles. (focus sensitive operator) (42)She even painted the garage. (focus sensitive operator) (43) Alan punched Owen and then <u>he</u> kicked <u>him</u>. (pronominal reference) Each sentence appeared at the end of a short paragraph, which embedded the sentence in a brief discourse context. We manipulated the information structure of each context to highlight

24

one of the competing interpretations. Based on previous research, we predicted that each interpretation should (or more weakly, could) also be accompanied by a particular prosodic pattern. For all items (with the exception of those with pronominal reference), the target sentence was followed by an additional sentence. The reason for this was that those sentences which seemed most naturally produced with a rising intonation seemed to end abruptly without such a continuation. For consistency across items, we included such a following sentence across all minimal pair members.

For the test items, two of the contexts varied the scopal relation between the quantifier and negation (quantifier > negation, negation > quantifier) in contexts that were designed to test whether each was accompanied by the predicted sentence-final contour (falling v. fall-rise, respectively). In one context, negation was associated with the presupposition and the QUD was therefore negative, while in a second context, the QUD was positive, and negation was associated with the assertion or focus. The second context also corresponds to Ward & Hirschberg (1985)'s type 1 uncertainty (whether to treat the quantifier as scalar).

We note here a key contrast with the two types of quantificational test items. With the sentences containing the universal quantifier *all* and negation, when *all* takes wide scope over negation, negation is at the VP level, and none of the discourse entities mentioned in the sentence have the property (i.e., none of the men went). When negation takes wide scope over the quantifier, negation is propositional, and the quantity of discourse entities mentioned in the sentence that have the property is not all, and possibly none. With the *many/most* and negation sentences, the situation is different. No matter what the scopal relation is (whether negation takes wide or narrow scope), the quantity does not vary: what varies is the *focus* on the quantity. For example, in (38) above, the number of alumni that Liam knows is always small. When *many*

takes wide scope over negation, emphasis is placed is on the number of alumni that Liam does not know (many). However, when negation takes wide scope over *many*, emphasis is placed on how many alumni he knows, which is few. See Baltazani (2002, 2003) for discussion.

In addition to the first two contexts, we presented these test items in one or two additional contexts, which allowed us to further evaluate how prosody can vary across discourse contexts, even when a scopal relation remains constant. For *many* and *most*, we created a third context in which we favored an interpretation where negation would take wide scope over *many* or *most*, but in which we predicted we might elicit a falling contour, since the target sentence appeared as an embedded clause. We modeled this item type directly after the example discussed by Kadmon & Roberts (1986).

For the universal quantifier items, we created two additional contexts that allowed us to achieve a fully crossed design for these items (negation in presupposition or assertion x prosodic contour). In the third context, we created a salient scalar alternative to the DP (e.g., *magnolias*), thereby inducing a fall-rise on this lexical item (Hirschberg, 1986; Ward & Hirschberg, 1985). In the fourth context, we manipulated the information structure to favor negation taking wide scope over the universal quantifier, as in the second context. However, instead of creating a scenario that gave rise to type 1 uncertainty and a fall-rise contour on *all*, we attempted to create a scenario that favored a falling contour. In (44), we present four contexts manipulating information structure for the test sentence *All the magnolias won't bloom*

(44) Four contexts for the target sentence *All the magnolias won't bloom*.

Context 1:

The township decided to plant magnolia saplings a number of years ago to line a path through the park. They have experienced lovely blossoms every year.

However, this year the area is experiencing less-than-standard rainfall, which means that they expect the magnolias to struggle this year, with only a few surviving. In fact, I think the situation is much more dire than that. <u>All the magnolias won't bloom</u>. They'll just have to wait till next year.

<u>Presupposition</u>: Some of the magnolias may not bloom. <u>QUD</u>: contains negation <u>Assertion</u>: None of them will.

<u>Scopal relation favored</u>: *all* > negation <u>Contour predicted to be favored</u>: falling <u>Context 2</u>:

A few years ago, the township decided to plant magnolia saplings to line a path through the park. The saplings on the north side were planted mainly in sand, and haven't been getting nearly enough nutrients. However, the soil near the south side is rich, and the magnolias are thriving there. <u>All the magnolias won't bloom</u>. But I bet the ones on the south side will.

<u>Presupposition</u>: All of the magnolias will bloom. <u>QUD</u>: does not contain negation <u>Assertion</u>: Some will.

<u>Scopal relation favored</u>: negation > *all* <u>Contour predicted to be favored</u>: fall-rise <u>Context 3</u>:

An aggressive beetle that targets magnolia trees is spreading through our area, and the magnolias are doomed. The township has been planning to take pictures for their website next month. The official photographer is concerned that there won't be beautiful rows of trees in the background for his pictures. I think he's worrying too much. <u>All the magnolias won't bloom</u>. However, there will still be other trees that will look just as lovely. <u>Presupposition</u>: Some of the trees will not bloom. <u>QUD</u>: contains negation <u>Assertion</u>: All of the magnolias will not bloom; other trees will.

<u>Scopal relation favored</u>: *all* > negation <u>Contour predicted to be favored</u>: fall-rise <u>Context 4</u>:

The weather recently has been conducive to plant growth, and all the trees are looking healthy. Some optimistic members of the township are predicting that each of the magnolia trees will give us lovely, fragrant blossoms to enjoy all season. But I think they're being rather unrealistic, and I keep telling them this. <u>All the magnolias won't bloom</u>. The odds of each of them blooming are pretty slim.

<u>Presupposition</u>: All the magnolias will bloom. <u>QUD</u>: does not contain negation <u>Assertion</u>: Not all of the magnolias will bloom.

<u>Scopal relation favored</u>: negation > *all* <u>Contour predicted to be favored</u>: falling

4.1.3. Procedure

Participants were recorded one at a time using an AT4040 Cardioid Capacitor microphone with a pop filter in a sound-attenuated recording booth, and amplified through an ART Digital MPA Gold microphone pre-amplifier. Stimuli were presented to participants using SuperLab stimulus presentation software (Cedrus Corporation, 2012) on a Macbook. For each trial, the participant first read the entire paragraph with the discourse context and target sentence silently. They then answered a comprehension question that tested for their understanding of the target sentence in the context. Finally, they read the entire paragraph out loud, this time recorded. The comprehension questions allowed us to filter out items for which participants did not answer the question correctly. Stimuli were divided into two blocks, so that members of a minimal pair were separated between blocks, and each block contained a token from each test or control item type. Test and control items were then pseudorandomized within each block.

Participants were run in two conditions. In the first condition, participants completed all 56 items within one session, which lasted approximately 45 minutes to an hour. They were told to read the items as naturally as possible. The second condition was constructed to alleviate some of the burden of the task, which we noted the first time around. Participants were presented with only one of the two blocks, and were again encouraged to read the items as naturally as possible, as though they were recording them for an audiobook or reading to children. They were also provided with an example of a written discourse beforehand (a section from a soap opera transcript) to model expressive reading. Later analysis revealed that the additional training did not elicit any difference in delivery, so no conditional analysis is presented.

4.1.4. Analysis

Target sentences were excised from the surrounding context using Praat speech analysis software (Boersma & Weenink, 2011) in order to conduct a series of analyses. They were then annotated with segments delineated from the onset and offset of each lexical item. For the controls items containing a *because* clause and negation or a focus sensitive operator, we collected the value of the maximum F0 on the relevant lexical items. For the focus sensitive operators, this was the head of the VP and DP that could host the focus (e.g., *painted* and *garage* in the sentence *She only painted the garage*). For the *because* sentences, this was the verb following negation and the final word in the sentence (e.g., *singing* and *audition* in the sentence *Georgia isn't singing because she's preparing for an audition*).

For sentences containing a *because* clause and negation, we predicted based on previous research that in the case where the *because* clause takes wide scope and negation is targeted at

the VP, there might be a prosodic break before the *because* clause (cf. Cooper & Paccia-Cooper, 1986; Koizumi, 2009). Therefore, in addition to the annotations based on lexical items, we also annotated 'pauses' for these items. However, because of the variability in the manifestation of the prosodic break within and among speakers, we could not systematically analyze this segment. Although some sentences had a clear pause inserted before *because*, others displayed a continuation rise on the previous lexical item. For those items where this lexical item did not end in an obstruent (e.g., *singing* v. *late*), this analysis became especially challenging to perform. We therefore did not analyze this aspect of the *because* sentences. We refer the reader to the abovementioned research for reported evidence of the regularity of such a prosodic feature for sentence interpretation.

For the test items, we conducted two main analyses. First, each file was coded for the type of sentence-final contour that was observed. To do this, two experimenters with musical training independently listened to each file (with all identifiers of context removed) and blindly coded it as either a falling or non-falling (fall-rise) contour. Rate of agreement was 80%. Any and all discrepancies were reconciled blindly afterwards with the assistance of a third blind coder, using as a comparison other clear exemplars from the participant whose items were in question. Second, we also conducted an acoustic analysis of the test items in order to look for any acoustic signatures in the speech signal that accompanied a difference in interpretation. In order to do this, Praat scripts were written and run on the annotated files to extract the relevant acoustic information. In the following section, we present each analysis in turn.

We excluded from analysis the following: tokens that had a comprehension score of 0 and tokens that had mis-starts, errors in pronunciation, use of the partitive in the quantificational phrase, glottalization in key lexical items, and/or errors in subject-verb agreement. Participants for whom there were not enough data points to perform an acoustic analysis across items after this filtering process were excluded from analysis (n=4). In addition, of the original contextsentence pairings, one set from universal quantifier *all* and negation items and one set from the *many/most* and negation items (the fourth item in each set in Appendix A) were excluded from analysis, since the comprehension scores for one or all of the items in each set were consistently at or below chance level across speakers, and speakers either reported difficulty in accessing the correct interpretations for these items, or experienced difficulty producing the items while being recorded. This conservative filtering process still left us with a rather large sample to work with across and test and control items and speakers, given our design.

4.1.5. Results

Here, we present the results of the analysis for each item type in turn, beginning with our control items and finishing with our test items involving quantification and negation.

4.1.5.1. Pronominal Reference

Perhaps surprisingly, participants did not mark pronominal reference with pitch accents. In the scenarios favoring a default, unstressed pronoun, participants produced the target pronoun in this manner 96.3% of the time. However, in the context supporting a stressed pronoun, with reversed reference, participants also produced the target pronoun as unstressed 95.6% of the time. This may be in part due to the fact that participants had much lower comprehension scores for the 'reversed' cases than for the 'default' cases (an average of 56.3% v. 90.6%), however, even for the most successful 'reversed pronominal reference' item in which comprehension scores were quite high (87.5%), participants were more likely than not to produce the pronoun as unstressed.

4.1.5.2. Focus-Sensitive Operators *Only* and *Even*

As an initial analysis of these items, we coded each item for the observed accenting pattern,

recording whether the main accent was on the head verb of the VP or on the DP for the *even* and *only* cases separately. We predicted that this pattern would be correlated with the focus pattern supported by the discourse context. The results are presented in Table **3**. Table 3: Pitch accenting patterns for the focus sensitive items *only* and *even*

		observed a	observed accenting	
_	focus	verb	DO	
even	VP	6	24	
_	DO	6	24	
only	VP	25	22	
-	DO	5	37	

Participants appeared to place the nuclear pitch accent on a syllable in the direct object for both sets of control items containing the focus sensitive operators. For *even*, there was absolutely no difference in accenting patterns for the individual items (p=1), regardless of the preceding context and the favored interpretation. For example, *she even painted the garage* was delivered similarly, regardless of the preceding context. This pattern held, despite the fact that participants' comprehension scores were the same for both context: 90.6% in each.

For *only*, the pattern was different. Participants' accenting pattern varied with the interpretation favored by the preceding context. As illustrated in Table **3**, a syllable in the direct object was more likely to be accented when the previous context favored a scalar contrast of the direct object (e.g., *Warren only likes the ORIOLES* [and not any other baseball team]). Given the results from *even*, however, this pattern could be taken as the default. Evidence for the preceding context playing a role in the pitch accenting pattern comes especially from the pattern exhibited when the preceding context favored a scalar contrast of the verb (e.g., *Warren only LIKES the Orioles*). With these items, participants were pulled away from this default pattern of placing the pitch accent on the direct object (Pearson χ^2 (1) = 16.92, *p*<.0001, ϕ = -.44), but were no more likely to place the accent on the verb than on the direct object (binomial probability *p*=.77). As

before, participants' comprehension scores were quite high: 81.3% for contexts favoring a verb contrast, and 100% for contexts favoring a contrast of the direct object.

Because the direct object was at the end of the declarative sentence, a comparison of the maximum F0 between the verb and the direct object proved difficult, as the F0 value continued to decline towards the sentence end. We therefore focused our attention on analysis of the location of the maximum F0 peak within the target items. To obtain this value, we calculated the maximum F0 within the target lexical item in the sentence, then calculated its location within the duration of that lexical item. This gave us a ratio that allowed us to measure the placement of F0 within the lexical item. These results are presented in Figure 1. We present this analysis of these control items, because the results demonstrate that participants produced regular and detectable acoustic differences correlated with contrasts in meaning during the experiment.

We performed a 2 x 2 ANOVA with word (V, DO) and context (focus on VP or focus on DO), comparing the maximum F0 location for the focus sensitive operator items. We excluded from analysis values that were more than 1.5 standard deviations away from the mean. We found a significant main effect of word (F(1, 181) = 12.2, MSE = .78, p=.0006), no main effect of context (F(1, 181) = .16, MSE = .01, p=.70), and a significant interaction (F(1, 181) = 4.38, MSE = .06, p=.038). Post-hoc Tukey's HSD tests revealed that the ratio of the location of Max F0 was larger (at the .01 level of significance) in the direct object in the context in which this focus pattern was predicted. While the difference was not significant for the verb, it approached significance in the right direction.

33



Figure 1: Maximum F0 location in sentences with a focus sensitive operator (*only, even*)

These results indicated that in the context in which the focus was predicted to be on the DO rather than the verb, there was a delayed placement of the maximum F0 in the direct object. One possible interpretation of these results is that the context set up a contrast between the salient scalar alternatives (i.e., the garage, as opposed to other things one could have painted; one mile, as opposed to more, and so on), thereby eliciting a fall-rise contour on these items.

4.1.5.3. Because clause and negation

For ambiguous sentences that contained a *because* clause interacting with negation, we also calculated the maximum F0 and the maximum F0 placement on two lexical items: the negated verb or predicate preceding the word *because* (e.g., *not <u>late because</u>, <i>isn't <u>singing because</u>*), and the final word in the sentence. We then compared these two values in both contexts. As with the focus sensitive operators, any possible difference in the second word between the two contexts was most likely washed out by its sentence-final position. However, the key value is the difference between the two words within a context. We present these results in Figure **2**.

Figure 2: Maximum F0 location in two words (word preceding *because* and sentence-final word) in sentences with a *because* clause and negation (e.g., *They're not* <u>late</u> because of his <u>driving</u>)



As in previous analyses, any values that were greater than 1.5 standard deviations away from the mean were excluded from analysis. A 2 x 2 (word x context) independent samples ANOVA revealed a main effect of word (F(1, 170) = 8.98, MSE = 19403.79, p=.003), no effect of context (F(1, 170) = 0, MSE = 2.32, p=1), and a significant interaction (F(1, 170) = 4.03, MSE = 8709.38, p=.046). Post-hoc Tukey's HSD tests revealed that there was a significantly higher maximum F0 on the first word than on the second word in the *because* > negation contexts at the .05 level of significance, and that the first word in this context was significantly higher than the first word in the negation > *because* context at the .05 significance. Such a pattern makes sense when we consider the interpretation of the 'because > negation' cases: the subject is NOT performing the action, and the reason for this state of affairs is being asserted (i.e., Georgia isn't SINGing, [and it's] because she's preparing for an audition.)

As with the focus sensitive operators, we also sought to compare the location of the maximum F0 in the key words in both contexts. See Figure **3**. Since the variance between the samples was unequal, we conducted two-tailed independent t tests assuming unequal variance, comparing the location of the maximum F0 in each of the key words between the two contexts. There was a significant difference in the Max F0 placement between the two contexts on the

second word (t(73)=-2.51, p=.01) – with the ratio larger (i.e., placement more delayed) in the negation > *because* context – but no significant difference in the first word for the two contexts (t(82)=-.13, p=.90).

Figure 3: Maximum F0 location in items with a *because* clause and negation (e.g., *They're not late because of his driving*)



Once more, intuitively, this difference in the second word makes sense. In the context in which negation takes wide scope, the subject IS performing the action, and what is asserted is that it is not for the reason made explicit that this is so (i.e., Georgia isn't singing because she's preparing for an au-DI-tion...She just likes to SING.) It has been claimed that these contexts should be more likely to exhibit a rising contour (cf. Koizumi, 2009). In addition, the two reasons are being contrasted, thereby eliciting a fall-rise contour in which the high pitch accent is expected to fall later in the word.

4.1.5.4. Quantification and negation

We not turn to our test items involving quantification and negation.

Analysis 1: Coding of Sentence-Final Contour

In our first analysis of these items, the excised sentences were coded blindly for falling versus non-falling contour. The percentage of observed falling contour for the sentences in each context is presented in Table **4**.
quantifier	context	scopal relation favored	% falling contour observed
all	1	all > negation	93.4
	2	negation > all	89.1
	3	<i>all</i> > negation	71.1
	4	negation > <i>all</i>	95.5
many/most	1	M > negation	91.3
	2	negation $> M$	65.1
	3	negation $> M$	63.0

Table 4: Predicted contour and observed falling contour for target sentences in each of the corresponding discourse contexts

A binomial probability analysis for the *all* sentences reveals that the frequency of observed falling contour is greater than chance for all contexts (contexts 1, 2, 4: p<.0001; context 3: p<.01). For the *many/most* sentences, the frequency of observed falling contour is greater than chance for context 1 (p<.0001), marginally significant for context 2 at p=.07, and not significant for context 3 (p=.10). A χ^2 analysis reveals a significant difference among the categories ($\chi^2(2) = 19.59, p$ <.0001). Thus, the sentence-final contour categorized as either falling or not falling is not at all a good indicator of scope for the *all*-negation sentences (*contra* any claims in the literature that it is a reliable indicator). However, in spite of the overall trend for a falling contour in the *many*-negation sentences, the difference in the sentence-final contour is much more likely than when negation takes wide scope. Thus, here we observe an item difference between the types of quantificational sentences.

Analysis 2: Acoustic Analysis

In our second analysis, we extracted acoustic information from each of the sentences, targeting two key lexical items: the quantifier (*all, many*, or *most*) and the sentence-final word. For the two types of test items, we conducted a two-way ANOVA on the quantifier and on the final word

comparing the factors of context and acoustic measure (treated separately): (a) maximum pitch (F0) in the word, (b) the location of the maximal F0 in the word (measured as a ratio over the entire word length, as described above), (c) the F0 standard deviation within the word, and (d) the duration of the word. We take each of these analyses in turn.

All and negation sentences

We begin with the results for the sentences containing *all* and negation, whose results are presented in Table **5**.

Table 5: Values of four key acoustic measures observed in each of the four discourse contexts for the quantifier *all* and the sentence-final word for sentences containing *all* and negation (e.g., <u>*All*</u> *the magnolias won't* <u>*bloom*</u>.) Significant differences are highlighted with dark border; marginally significant differences indicated with a dashed line.

quantifier: all				sentence-final word				
Acoustic measure	Context	Mean	Std. Error	Acoustic measure	Context	Mean	Std. Error	
maximum F0	1 ∀>¬	230.90	15.13	maximum F0 (Hz)	1	192.56	85.67	
(Hz)	2 ¬>∀	245.23	15.13		2	162.84	51.40	
	3 ∀>¬	247.30	15.13		3	181.73	39.85	
	4 ¬>∀	230.78	16.11		4	190.17	47.39	
maximum F0	1 ∀>¬	.76	.05	maximum F0 location (ratio)	1	.38	.29	
location (ratio)	2 ¬>∀	.79	.05		2	.44	.34	
	3 ∀>¬	.79	.05		3	.41	.25	
	4 ¬>∀	.89	.05		4	.48	.29	
F0	1 ∀>¬	26.06	3.94	FO	1	29.45	17.07	
standard deviation (Hz)	2 ¬>∀	28.73	3.94	standard deviation (Hz)	2	21.01	14.36	
()	3 ∀>¬	25.83	3.94		3	28.43	13.13	
	4 ¬>∀	28.40	4.19		4	30.43	17.25	
word	1 ∀>¬	170	10	word duration (ms)	1	391	65	
duration (ms)	2 ¬>∀	160	10		2	391	65	
	3 ∀>¬	170	10		3	448	59	
	4 ¬>∀	168	11		4	349	75	

Turning first to the analysis *all*, we found no main effects when directing our attention to the quantifier itself, which, recall, was in sentence-initial position: maximum F0: F(3, 66) = .34, *p*=.80; maximum F0 location: F(3, 66) = 1.09, *p*=.36; F0 standard deviation: F(3, 66) = .15, *p*=.93; word duration: F(3, 66) = .19, *p*=.90. Planned pairwise comparisons between each of the contexts within each of the four acoustic measures supported this lack of significance, although

there was a marginally significant difference between contexts 1 and 4 for the maximum F0 location (p=.09): the peak F0 for context 4 (where negation > *all*) appeared slightly later than the peak F0 for context 1 (where *all* > negation). Note that this trend is consistent with the results for the *because*-negation sentences, when negation took wide scope.

Turning to the final word for sentences containing *all* and negation, we found a similar pattern of null results for the first three acoustic measures: maximum F0: F(3, 66) = .87, p = .46; maximum F0 location: F(3, 66) = .31, p=.82; F0 standard deviation: F(3, 66) = 1.27, p=.29. However, there was a significant main effect of word duration: F(3, 66) = 6.20, p=.001. Planned pairwise comparisons between each of the contexts within each of the acoustic measures confirmed the lack of significance among the contexts for the first three acoustic measures, although the difference between contexts 2 and 4 for the F0 standard deviation was marginally significant (p=.09). The significant main effect of final word duration was driven by a highly significant difference between contexts 3 and 4 (p = <.0001), and a significant difference between contexts 1 and 3 and 2 and 3 (both p=.01). There was a marginally significant difference in the final word duration between context 1 and 4 and between context 2 and 4 (both p=.08), and no significant difference in the final word duration between contexts 1 and 2 (p=1.0). Thus, duration of the final word was longest in context 3, where the scopal relation favored was all > negation, but there was a contrast in scalar alternatives (e.g., magnolias v. other trees), predicting a fall-rise contour would be exhibited. And the duration was shortest in context 4, where the scopal relation favored was negation > *all*, and the predicted contour was falling.

Many/Most and negation sentences

We turn next to the results for the sentences containing *many/most* and negation, whose

results are presented in Table 6.⁴

Table 6: Values of four key acoustic measures observed in each of the three discourse contexts for the quantifiers *many* and *most* and the sentence-final word for sentences containing *many/most* and negation (e.g., *Liam doesn't know <u>many alumni</u>*.) Significant differences are highlighted with dark border.

quantifier: many or most				sentence-final word			
Acoustic measure	Context	Mean	Std.	Acoustic measure	Context	Mean	Std.
			Error				Error
maximum F0	1 M>¬	186.54	38.64	maximum F0	1	189.97	72.32
(Hz)	2 ¬>M	195.40	42.32	(Hz)	2	189.15	39.36
	3 ¬>M	187.19	44.78		3	167.32	40.57
maximum F0	1 M >¬	.35	.30	maximum F0	1	.27	.17
location (ratio)	2 ¬>M	.43	.26	location (ratio)	2	.26	.21
	3 ¬>M	.43	.23		3	.34	.30
F0	1 M >¬	10.68	6.21	F0	1	24.79	14.70
standard deviation (Hz)	2 ¬>M	10.59	7.25	standard	2	25.33	10.88
	3 ¬>M	13.69	11.59	deviation (Hz)	3	21.72	14.68
word	1 M >¬	277	24	word	1	520	51
duration (ms)	2 ¬>M	308	38	duration (ms)	2	462	51
	3 ¬>M	300	34		3	478	69

The ANOVA run on the quantifier *many/most* uncovered no main effects for three of the measures: maximum F0: F(2, 44) = .20, p=.82; maximum F0 location: F(2, 44) = .48, p=.62; F0 standard deviation: F(2,44) = .61, p=.55. However, there was a significant main effect of word duration: F(2, 44) = 3.68, p=.03. Planned pairwise comparisons between each of the contexts within each of the four acoustic measures complemented these findings to reveal no significant difference among the three contexts for the first three acoustic measures, but significant

⁴ We collapse over *many* and *most* in our analysis, but note that we are never comparing *many* to *most*.

differences among the contexts for word duration. While contexts 2 and 3 were not significantly different from each other (p=.52), there was a significant difference between contexts 1 and 2 (p=.01) and a marginally significant difference between contexts 1 and 3 (p=.06).

Finally, we turn to the ANOVA on the final word of sentences with *many/most* and negation. There was no main effect for the first three acoustic measures: maximum F0: F(2, 44) = .86, p=.43; maximum F0 location: F(2, 44) = .46, p=.64; F0 standard deviation: F(2, 44) = .30, p=.74. However, as before, there was a significant main effect of word duration: F(2, 44) = 4.03, p=.03. Planned pairwise comparisons between each of the contexts within each of the four acoustic measures revealed no significant differences between any of the contexts for the first three acoustic measures, but significant differences between the three contexts for final word duration: context 1 v. 2 p=.009, 1 v. 3 p=.05, but contexts 2 v. 3 p=.45. Combining these two analyses for *many/most*, we find that the length of the quantifier was shortest in context 1, with the duration in contexts 2 and 3 not differing from each other. At the same time, the duration of the final word was longest in context 1, with the duration in contexts 2 and 3 not differing from each other.

Focusing only on the final word, we observe an interesting comparison between the *all* and the *many/most* cases: the longest final word duration for each set of target sentences was observed in a context in which the quantifier > negation scopal relation was favored (context 3 for *all* and context 1 for *many/most*). The duration was shortest for the *all* sentences in a context favoring a negation > quantifier relation (context 4). These durations do not, however, entirely correlate with the predicted or observed contour, since all four contexts of the *all* sentences exhibited a robust falling contour, but the first context for the *many/most* sentences was the contour that had the highest percentage of falling contour. Thus, one could try to argue that a

falling contour is correlated with a longer word-final duration, but we are left trying to account for why context 1 for the *all* sentences displayed no such difference from contexts 2 and 4.

Likewise, one could try to argue that a quantifier > negation scopal relation consistently displays an acoustic correlate, while a negation > quantifier scopal relation consistently displays a diverging pattern. There may, in fact, be reason to make a case for this distinction with the *many/most* sentences, but there appears to be no such evidence for this distinction in the *all* sentences, which were all invariably magnets for a falling contour.

4.1.6. Results

The findings presented in the preceding section lead us to the following conclusions. First, we replicate and extend previous conclusions for sentences with the focus-sensitive operators *only* and *even* and sentences with *because* and negation by demonstrating that speakers reliably vary their production of such sentences according to scopal relation and intended interpretation favored by the discourse context. Our analysis of sentences containing the focus-sensitive operators *only* and *even* revealed a general trend to place a pitch accent on the direct object, rather than the verb. However, with sentences containing *only*, we observed a tug away from this pattern towards accenting the verb when the focus was placed at the VP level. In addition, we observed a strong overall trend for participants to delay the placement of the maximum F0 in the direct object in precisely that context in which the focus was predicted to be on the DO rather than the verb.

When we analyzed the sentences containing *because* and negation, we found a significantly higher maximum pitch on the first word than in the second word when *because* took scope over negation, and this maximum pitch on the first word in these contexts was also higher than the maximum pitch in the first word of the negation > *because* contexts. No such difference

in the maximum pitch was found between the two words in the negation > *because* contexts. We also observed that in the contexts in which negation takes wide scope over the *because* clause, denying the reason for the subject of the sentence performing the action in question, the pitch accent on the second word comes later than in the *because* > negation context.

Our second main conclusion is that while speakers do produce ambiguous sentences in a way that could disambiguate them for the hearer with many control items, the results are less clear-cut for sentences with quantification and negation. Contrary to previous statements in the theoretical literature that speakers reliably signal their interpretation with the sentence-final contour, we found no such pattern for sentences with the universal quantifier *all* and negation. And while we found a difference along those lines with the *many/most* sentences, in that a falling contour was far more likely with sentences in which negation took narrow scope, speakers more often than not exhibited a falling contour, and did not robustly signal negation taking wide scope with a fall-rise contour. Instead, the differences appear to occur at a lower level—in the form of the placement of the maximum pitch within a word, but most often with a difference in word duration. However, even here, such acoustic correlates are hard to pin down as a definite signal to interpretation. For example, as described above, duration of the final word is longest when negation takes narrow scope, but this is neither a necessary nor sufficient cue to this scopal relation.

One possibility, however, is that this omnibus analysis of the data may mask intra- and interspeaker variability in the production of these sentences. That is, speakers may be recruiting a variety of cues (e.g., word duration, maximum pitch, maximum pitch placement, and so on) differentially from utterance to utterance, and/or each speaker may have his/her own strategy for disambiguation. It is not possible to tease apart this possibility from the set of findings at this

point, without an extensive item by item and individual by individual analysis. We leave that research project open as an avenue for future researchers to follow. For now, we leave it as open possibility that individual speakers are actually better at consistently identifying strategies for signaling interpretation with prosodic and acoustic cues than the current work (averaging over speakers) suggests. We present our contribution as showing that sentence-final contour does not appear to be *the* indicator of interpretation, and that there are other cues, such as duration, which may be informative.

The findings from our production study raise an important question, however: even if speakers do reliably use surface-level cues to signal their intended interpretation, can hearers then recruit these cues to arrive at the interpretation that was intended by the speaker? To date, such evidence for quantificational sentences in English has remained elusive in small-scale studies aimed at addressing this question. Moreover, we are not aware of any extensive work on sentence perception that has systematically manipulated the requisite discourse-relevant and experimental variables to tackle this question. In the next two sections, we present two complementary perception studies, which demonstrate that in the best case scenario – the scenario in which speakers reliably produce two distinct versions of the same exact sentence to signal two scopally different relations (in a manner consistent with previous claims concerning this production) – hearers can arrive at the correct interpretation.

5. Experiment 2: Perception Study 1

The joint purpose of the two perception studies was to determine whether hearers can use surface level prosodic and acoustic cues provided by speakers to arrive at the intended interpretation of scopally ambiguous sentences. To this end, each of the perception studies served a specific purpose. In Perception Study 1, we sought to determine whether hearers could use the prosodic

45

delivery of a target sentence to properly situate it in a discourse context indicated by two possible continuation sentences. In Perception Study 2, we sought to determine whether hearers could use the thread of a discourse context to choose the prosodic version of a target sentence that was most compatible with the interpretation favored by that context.

Tokens for both perception studies were contributed by a combination of naïve speakers from the production study reported above (three in Perception Experiment 1 and two in Perception Experiment 2) and an experimenter (one of the authors, a female who has had extensive vocal musical training and is well acquainted with the theoretical and experimental literature – and was therefore well acquainted with the prosodic patterns claimed to correlated with the various interpretations). The two experiments were run completely separately. Each perception experiment began with a brief training session in which participants were introduced to the task structure, and the possibility of disambiguating sentences with prosody, using items with ambiguous pronominal reference. Each experiment was completed in under 30 minutes, with Perception Experiment 2 being the shorter of the two.

5.1. Experimental Method

5.1.1. Participants

49 undergraduates participated in Perception Experiment 2. Data from four participants were excluded from the analysis, because the participants indicated that English was not their native language. In addition, data from one participant were excluded due to consistently low response times (many below 150 ms).

5.1.2. Stimuli

Items for Perception Experiment 1 were contributed by four different speakers: three sets from naïve speakers taken from the production study (one male and two females) and one set

produced by the experimenter. The naïve speakers were selected based on their high comprehension scores in the production experiment (above 75%) and their consistent and clear production of distinct versions of the two interpretations of each sentence (based on contexts 1 and 2), which largely reflected the manner of production discussed in the literature.

There were 48 experimental items. These included 24 minimal pairs of sentences in which the same sentence was produced in two distinct manners, each favoring an interpretation supported by a previous discourse context. Each speaker contributed six minimal pairs: two with *all* and negation, one with *many/most* and negation, two with *because* and negation, and one with one of the focus-sensitive operators (*only* or *even*). These items are noted in the Appendix. The minimal pairs for each speaker were blocked by speaker, and then further divided within the block to separate the minimal pair members. During the experiment, the presentation of the items within the blocks was randomized by the stimuli-presentation software. The order of the blocks was predetermined: the female experimenter, a naïve male, a naïve female, and a second naïve female – one who had a very high comprehension score in the production study (92.9% total comprehension correct).

5.1.3. Procedure

Items for both of the perception experiments were presented using the Superlab stimulus presentation software (Cedrus Corporation, 2012) with headphones at iMac computers in a quiet laboratory setting. Each item in Perception Experiment 1 had the following structure. First, participants viewed the target sentence in the middle of the screen (e.g., *All the moms didn't allow eyeliner*). This visual stimulus was accompanied by a speaker's production of the sentence, which was intended to be uttered as part of a cohesive discourse context that favored one possible interpretation. Pitch tracks for two renditions of one target by the experimenter (speaker

1) are presented in Figure **4**. Additional examples for both perception experiments are presented in Appendix B.

Figure 4: Pitch tracks for two renditions of the target sentence *All the moms didn't allow eyeliner* by speaker 1 (the experimenter) from context 1, *all* > negation (left) and context 2, negation > *all* (right)



The auditory stimulus was repeated three times. The participant then saw a second screen, on which were two possible continuations to the target sentence, one above the other (e.g., A: *They were all in agreement.*; B: *Only the moms of the older girls let their daughters wear it.*).

Participants were asked to choose the most natural continuation. In this way, we placed the target sentences into a mini discourse. The correct choice was counterbalanced between A and B. We coded a correct response as one that corresponded to the interpretation intended by the speaker based on the production study. Participants had a maximum of 15 ms to make their decision. If no response was made during that time, the trial would end, and the next one would begin. There were eight instances where a participant made no response during the time allotted. Responses registered in less than 200 ms (n=4) were not included in the analysis.

5.1.4. Results

The dependent measure was the percentage of correct responses for each item type, overall and

by scopal relation (negation taking wide or narrow scope relative to the target lexical item). The overall results for those items involving negation are presented in Table 7.

Itom Tuno	% correct (SE)	speaker 1	speaker 2	speaker 3	speaker 4
item Type		(exp'r, F)	(382, M)	(369, F)	(362, F)
total	67.8 (.02)	73.1	62.3	66.7	69.1
all	65.2 (.02)	72.2	62.5	68.8	56.3
<i>all</i> > negation	63.9 (.03)	61.4	67.1	65.9	60.2
negation > all	66.4 (.03)	83.0	58.0	71.6	52.3
many/most	69.3 (.02)	77.3	64.8	55.7	79.6
M > negation	62.5 (.04)	77.3	54.6	29.6^{5}	88.6
negation $> M$	76.1 (.03)	77.3	75.0	81.8	70.5
because	69.3 (.02)	69.9	59.7	75.6	71.6
<i>because</i> > negation	62.4 (.03)	61.4	52.3	71.6	63.6
negation > because	76.0 (.03)	78.4	67.1	79.6	79.3

Table 7: Average percentage correct overall per item type and per speaker

We conducted two main analyses of the data. First, we sought to determine if the response pattern was significantly different from chance. We first ran a binomial probability test on the overall average (far left column) to determine whether the average scores for each item type and each corresponding scopal relation was above chance (p=.5, 99% CI). Indeed, all averages were significantly above chance level: *all* > negation: 2.7, *p*<.01; negation > *all*: 3.1, *p*=.002;

⁵ This egregious departure from all other averages appears to have been the result of hearers overwhelmingly interpreting speaker 3 (369)'s production of this item as having the reverse scope. Upon further inspection of this file, we discovered that for some reason, this speaker appeared to place contrastive focus on the word *enjoy* in the sentence *Neil doesn't enjoy most musicals*. This prosodic pattern may have led hearers to interpret the contour as favoring the reverse scope. For context 2 (negation > M), this speaker uttered the sentence as expected, with a clear fall-rise contour on *most* and at the end of the sentence. Hearers exhibited correct responses to that item over 80% of the time. many/most > negation: 2.5, p=.01; negation > many/most: 5.1, p<.0001; because > negation: 2.3, p=.02; negation > because: 5.1, p<.0001. This is the first indication that hearers were able to successfully use the speakers' prosodic delivery of the target sentences to arrive at the correct interpretation.

This analysis was complemented by a χ^2 analysis on the distribution of responses, based on the number of participants responding at each percentage level. As would be expected, the findings were significantly above chance: *all* > negation: $\chi^2(7) = 20$, *p*<.01, negation > *all*: $\chi^2(7)$ = 32, *p*<.01; *many/most* > negation: $\chi^2(4) = 21.45$, *p*<.01; negation > *many/most*: $\chi^2(4) = 31.45$, p < .01; *because* > negation: $\chi^2(7) = 19.27$, p <.01; negation > *because*: $\chi^2(7) = 33.09$, *p*<.01. The histograms for each item type are presented in Figure **5**, comparing the distribution for each scopal relation against a normal distribution.

Figure 5: Distribution of responses for three item types and two scopal relations within each for Perception Experiment 1



Second, we sought to determine whether the response pattern varied across the factors we manipulated in the experiment. We therefore ran a 4 x 2 x 3 ANOVA comparing speaker (each of the four speakers), scope of negation (wide, narrow), and lexical item interacting with negation (*all, many/most, because*). This analysis revealed a main effect of speaker (F(3, 1056) = 3.94, *p*=.008), a highly significant main effect of the scope of negation (F(1, 1056) = 19.42, *p*<.0001), and significant interactions between speaker*scope of negation (F(3, 1056) = 5.35,

p=.001), speaker*lexical item (F(6, 1056) = 5.10, p<.0001) and speaker*scope of negation*lexical item (F(6, 1056) = 6.44, p<.0001. There was a marginally significant effect of scope of negation*lexical item (F(2, 1056) = 2.70, p=.07). There was no main effect of lexical item (F(2, 1056) = 1.61, p=.20). Planned pairwise comparisons confirmed no significant differences among lexical items, and a highly significant overall differences between the two scopes of negation (p<.0001), with better performance when negation took wide scope. There were also significant differences between the speakers: speaker 2 (382), who had the lowest average overall, was significantly different from speaker 1 (the experimenter) (p=.001) and speaker 4 (362) (p=.04) (the speakers with the two highest averages). Speaker 1 (experimenter) and speaker 3 (369) were also significantly different from each other (p=.05).

Following the experimental session, participants were asked to comment on their strategy for responding to the items, and whether any aspect of the items themselves made making a decision either easier or harder (following an approach used successfully by Syrett *et al.*, 2011). We took these open-ended responses and categorized them based on key words in each participant's response. Responses were coded as mentioning 'word-level stress' if they included any of the following words: *inflection* (with regards to specific words), *tone(s)* (on words), *emphasis, stress, loud,* or *soft.* Responses were coded as mentioning 'phrase-level intonation' if they included words such as *up, down, end, rais(ed), low(er), drone, monotone, pitch,* and *inflection* (with regards to entire sentence). Responses were coded as mentioning 'prosodic break' if they contained words such as *breath(e), pause(s),* or *break(s).* Only three of the participants' responses could not be coded in this manner, because their responses were too vague. Two participants failed to complete a survey. Note that the nature of this coding method meant that participants' responses could be coded as belonging to one or more categories. We

present the distribution of responses in Table 8.

Table 8: Participants' categorized responses to the post-experiment survey in PerceptionExperiment 1 regarding their response strategies

Category	N participants	% Participants
Word-level stress	33/42	78.5
Phrase-level intonation	12/42	28.6
Prosodic break	3/42	7.1

5.1.5. Discussion

The results from this perception experiment demonstrate that when speakers use prosody to favor an interpretation of a scopally ambiguous sentence that was supported by a discourse context, hearers can successfully recruit this information to assign the correct interpretation. The variability in the averages among speakers and items further illustrates the importance of examining a range of lexical items, contexts, and speakers. Indeed, it would not be surprising to find further effects of age, dialect, or other factors relevant to this phenomenon. Further, not only were hearers able to use the auditory information encoded in the production to make their selection, they seemed to be sensitive to precisely those aspects that speakers were manipulating to favor an interpretation, as evidenced by their responses to the post-experiment survey, presented in Table 8. This may not seem so surprising, since the only way that participants could have systematically made their decision in the experiment was to read and listen to the target sentences. But on a more fine-grained level, participants' responses indicate that they were attending to specific aspects of the prosodic delivery to make their choice, and did so successfully. We also note that most hearers did not indicate that they attended to whether the sentence was rising or falling, suggesting that there were other cues beyond sentence-final contour that enabled them to make their decision.

6. Experiment 3: Perception Study 2

Our aim in Perception Experiment 2 was to determine whether participants could use the information structure provided by the context to identify the version of the sentence that was a best fit, given the interpretation that had been favored.

6.1. Experimental Method

6.1.1. Participants

37 undergraduates participated in Perception Experiment 2. 10 additional non-native students participated, but their data were not analyzed.

6.1.2. Stimuli

The auditory stimuli were provided by three of the speakers from the previous Perception Experiment (the experimenter and the two naïve females). The male was excluded, because participants from the previous experiment occasionally reported difficulty with his files in the post-experiment survey. Indeed, his pitch range was not as wide, and his overall response scores were the lowest of the four. Each speaker provided six minimal pairs of sentences: two with *all* and negation, one with *many* and negation, two with *because* and negation, and one with a focussensitive operator (*only* or *even*). This resulted in 18 minimal pairs, for a total of 36 sentences per participant. Each sentence was part of an item that consisted of a brief discourse context shown on screen, followed by a forced choice of two versions of the target sentence in the context.

6.1.3. Procedure

As in the previous experiment, the experimental session began with a brief training session in which participants became acquainted with the experimental setup. The items in the training session involved disambiguation with pronominal reference. Participants then proceeded on to the test session. Each item had the following structure. First, participants were shown a discourse

context, which they progressed through line by line in a self-paced cumulative window fashion. An example is provided in 0. The target sentence was indicated with << >>.

(45) Context favoring the *all* > negation interpretation of *All the moms didn't allow eyeliner*.

Several young girls wanted to have a make-up party together. Some of them thought their mothers wouldn't let them use eyeliner. In fact, the moms were all on the same page. <<All the moms didn't allow eyeliner.>>

The girls were limited to mascara and blush.

Next, participants were presented with two auditory versions of the target sentence sequentially and each accompanied by 'A' or 'B' on the screen. A was always a falling contour, and B a nonfalling contour. A third screen then appeared with A and B on it, and participants made their selection of the best match, given the preceding discourse context.

6.1.4. Results

As in the previous experiment, the dependent measure was the percentage of correct responses for each item type, overall and by scopal relation (negation taking wide or narrow scope relative to the target lexical item). Responses registered after 3.5 seconds were not included in the analysis. We used this metric, because the unlike the previous experiment, where participants were asked to choose between two sentences, in this experiment, they were asked to make a decision between two sequential sounds. A prolonged delay might have had the effect of minimizing whatever features of each sound file and contrast between the two items participants were holding in their working memory. We also piloted and verified that 3.5 seconds was an ample amount of time to make the choice.

The results are presented in Table 9. Immediately apparent is the fact that hearers were

quite successful at identifying the prosodic match, given the preceding discourse context. Responses were well over 50% in all cases, with the exception of the 'negation > all' items, which we return to momentarily.

Table 9: Average percentage correct overall per item type and per speaker in PerceptionExperiment 2

Itom Tuno	% correct (SE)	speaker 1	speaker 3	speaker 4
item Type		(exp'r, F)	(369, F)	(362, F)
total	78.1 (.02)	76.1	81.0	76.8
all	65.0 (.02)	71.2	67.0	56.8
all > negation	76.9 (.03)	82.9	75.7	72.2
negation > all	53.1 (.03)	59.5	58.3	41.4
many/most	86.5 (.02)	77.8	90.1	91.8
M > negation	88.8 (.03)	88.9	85.7	91.7
negation $> M$	84.3 (.03)	88.9	94.4	91.9
because	82.4 (.02)	79.5	85.8	81.8
<i>because</i> > negation	79.3 (.03)	77.0	85.1	75.7
negation > because	85.4 (.03)	81.9	86.5	87.8

As in the previous experiment, we ran three analyses, the first two evaluating difference from chance, and the third an ANOVA investigating effects of the factors we manipulated in the experiment.

A binomial probability test (p=.5, 99% CI) on the overall average for each item type and each corresponding scopal relation revealed that nearly all averages were significantly above chance level: *all* > negation: 5.0, p<.00001; *many/most* > negation: 27.4, p<.00001; negation > *many/most*: 6.5, p<.00001; *because* > negation: 7.2, p<.00001; negation > *because*: 8.8, p<.00001. The one chance-level pattern was from negation > *all*: -1.0, p=.31. Thus, in the majority of cases, hearers were overwhelmingly successful at pairing the information from the discourse context with the prosodic version of the target sentence that best matched the favored interpretation and scopal relation. A χ^2 analysis on the distribution of responses by participant revealed that it was significantly different from chance: *all* > negation: $\chi^2(6) = 44.5$, *p*<.0001, negation > *all*: $\chi^2(6) = 23.33$, *p*=.0007; *many/most* > negation: $\chi^2(3) = 56.51$, *p*<.0001; negation > *many/most*: $\chi^2(3) = 43.54$, p < .0001; *because* > negation: $\chi^2(6) = 46.39$, p < .0001; negation > *because*: $\chi^2(6) = 52.44$, *p*<.0001.

Finally, we ran a 3 x 2 x 3 ANOVA comparing speaker (each of the three speakers), scope of negation (wide, narrow), and lexical item interacting with negation (all, many/most, *because*). This analysis revealed a highly significant main effect of lexical item (F(2, 650) = 25.7, p < .0001) and a significant main effect of the scope of negation (F(1, 650) = 19.42, p < .0001). These were accompanied by a significant scope of negation*lexical item interaction (F(2, 650) = 11.52, p < .0001). This time, without speaker 2 (382) from the previous production experiment, there was no main effect of speaker (F(2, 650) = 1.36, p=.26), and no effect of speaker*scope of negation (F(2, 650) = 1.59, p=.21). There was, however, a significant speaker*lexical item interaction (F(4, 650) = 3.33, p=.01), and a marginally significant interaction of speaker*scope of negation*lexical item (F(4, 650) = 1.92, p=.11. Planned pairwise comparisons revealed significant differences between *all* and each of the other two lexical items (p < .0001), but no difference between *because* and *many/most* (p = .19). There was no difference among any of the speakers. There was a significant difference between negation taking narrow v. wide scope (p=.005), driven by higher scores when negation took narrow scope overall. (However, this did not hold for all items, as evidence in Table 9.)

6.1.5. Discussion

The findings of Perception Experiment 2 combined with those from Perception Experiment 1 demonstrate without a doubt that hearers can indeed use prosodic information in order to disambiguate scopally ambiguous sentences and arrive at the interpretation intended by the

speaker. In this experiment, participants recruited the information structure from a discourse thread to assign an interpretation to a sentence, then selected the prosodic rendition of this sentence that best corresponded to that interpretation. Hearers in this experiment were very successful in this task, with the notable exception of the *all* sentences, particularly where negation took wide scope. It is interesting to note that in these cases, hearers were largely at chance. Thus, they were not displaying a bias towards one reading or another.

It cannot be that all sentences with negation taking wide scope caused them difficulty, as they were quite successful with the *many/most* cases and the *because* cases. Indeed, in these cases, hearers were highly successful with both scopal relations. What these latter two sentence types have in common – and where they contrast with the all sentences – is that in these two sentence types, negation precedes the quantifier *many/most* or *because*. As a consequence, we speculate that a few things may be at play. First, when negation precedes the lexical item in question (either the quantifier or *because*), this surface structure may make it easier for hearers to access an interpretation where negation takes wide scope. This may be because it allows them to more easily access sentential negation (as opposed to VP-level negation), and/or because it makes it easier to focus negation (in response to a positive QUD and a presupposition without negation). Second, when negation precedes a quantifier in object position, the quantifier may be able to host more informative prosodic information than when it is in subject position. Finally, when a hearer is processing the sentence incrementally, they may have accumulated enough relevant information in parsing the sentence before they hit the quantifier or because later in the sentence that they are in a better position to integrate the information and access the correct interpretation. Future research aimed at identifying the source of the difference between the *all* and *many/most* sentences would add to our understanding of how prosody is recruited and

implemented incrementally, and what the implications are for a quantifier appearing in different syntactic positions.

7. Conclusions and Discussion

This work is, to our knowledge, the first to systematically test seminal claims by Jackendoff (1972) that differences in scopal relations in sentences with quantification and negation in English are tied to differences in prosodic contours for both speakers and hearers, to probe for acoustic correlates elsewhere in the sentence beyond intonational contour, and to control for elements of the discourse context that would highlight a connection between information structure, prosody, and scope. Taking stock across our production and perception experiments, we found the following. Consistent with previous experimental studies, we found that speakers did use prosody (loosely construed) to disambiguate control sentences involving the focussensitive operators *only* and *even*, and sentences with *because* clauses interacting with negation. These differences manifested themselves the maximum F0 and the placement of this peak within the focused item. Thus, we know that our speakers were providing cues to disambiguation for sentences other than our test sentences with quantification and negation. For these test sentences, we found differences among quantificational items. While sentences with *all* in subject position were consistently produced with a falling contour, sentences with *many* or *most* were less likely to be produced with a falling contour when negation took wide scope - a pattern predicted by previous theoretical research (e.g., Kadmon & Roberts, 1986).

In spite of considerable variability within and among speakers in their production of these sentences, leading to a lack of significance along a number of acoustic measures, we did observe significant differences correlated with scopal relations. Here, too, however, we also observed item differences in acoustic information. For the *all* sentences, the ratio of the maximum F0

placement was greatest on the quantifier in a context where negation took wide scope. For the *many/most* sentences, duration of *many* or *most* was shortest when *many* or *most* took wide scope over negation. For both types of sentences, though, the longest final word duration was observed in a context in which the quantifier > negation scopal relation was favored (context 3 for *all* and context 1 for *many/most*). Thus, we found that sentence-final contour does not appear to be *the* indicator of interpretation, and that there are other acoustic cues, such as word duration, which may be informative. However, these low-level acoustic correlates may hard to pin down as a definite signal to interpretation, given the observed variability among contexts and speakers.

We then took a subset of the production files, supplemented by an experimenterproduced set, in which speakers regularly disambiguated among minimal pairs of sentences by manipulating prosody, and asked whether – given these clear prosodic differences correlated with sentence-level meaning – hearers could use such information in the speech signal to arrive upon correct, intended interpretation. In two sets of perception experiments, we provided an affirmative answer to this question. We note that in the perception studies, too, we found variability in the response rates for the speakers and for the items, complementing the production study and illustrating the importance of examining a range of lexical items, contexts, and speakers.

In Perception Study 1, we investigated whether hearers could use prosody to situate a target sentences in a discourse context that was indicated by two possible continuation sentences. We found that when speakers use prosody to favor an interpretation of a scopally ambiguous sentence that was supported by a discourse context, hearers can successfully recruit this information to assign the correct interpretation. A post-experiment survey also indicated that hearers were attending to prosodic aspects when making their choice. In Perception Study 2, we

59

investigated whether hearers could use the thread of a brief discourse context to choose the prosodic version of a target sentence that was most compatible with the interpretation favored by that context. Here, we found that participants recruited the information structure to assign an interpretation to a sentence, and were able to select the prosodic version of the sentence that best corresponded to that interpretation. The only notable exception to participants' overwhelmingly successful performance in this task was with universal quantification-negation sentences where negation took wide scope. This outlier underscores the importance of testing a range of quantifiers and syntactic positions, as generalizations cannot be made just from the universal quantifier *all* in subject position alone.

Thus, we conclude that *speakers can (although not uniformly) and hearers do* recruit prosodic and acoustic cues to disambiguate scopally ambiguous sentences precisely in those cases where these auditory cues are informative about the information structure in the discourse context, which is in turn informative about scopal relations. Given these findings, we must also conclude that psycholinguistic studies focused on participants' ability to disambiguate scopally ambiguous sentences (perhaps even more specifically, with negation, or perhaps more generally involving the interaction between two logical operators) should take auditory cues in the speech signal into account. One area in which this conclusion is directly relevant is in acquisition research aimed at identifying whether language learners are able to access the full range of interpretations of quantifier-negation sentences, such as our test sentences and the prototypical sentence from this line of research in (46).

(46) Every horse didn't jump over the fence.

Musolino (1998)'s initial observation from his dissertation work was that children for whatever reason seem to be locked into the reading where negation takes narrow scope (i.e., the

60

'surface scope' reading) and have great difficulty accessing the 'inverse scope' reading where negation takes wide scope (his 'observation of isomorphism'). This observation has been replicated across a variety of languages by a number of different researchers, using lexical items beyond the universal quantifier in subject position (cf. Musolino, 2011 for a review). However, subsequent research manipulating a variety of contextual factors has revealed that children are in fact able to access the inverse scope reading above chance level, and that it is indeed within their grammatical capacity (cf. Gualmini *et al.*, 2008; Musolino & Lidz, 2006; Viau, Lidz, & Musolino, 2010). One might ask, given the current results, what role prosody has to play in children's performance. Musolino & Lidz (2006) remark in their footnote no. 13 that "we are not aware of a single study on the acquisition of universal quantification in which prosodic cues are manipulated." To our knowledge, at the time of writing this paper, this appears to still be the case.

Previous authors indicate in passing that they were sensitive to the possible role of prosody, but never systematically controlled for this factor as an independent variable. For example, Lidz & Musolino (2002) say that "When making these statements, the experimenter holding the puppet was instructed to say the sentences in a way that is the most naturally compatible with the appropriate reading on which the sentence was a true description of what had happened in the story. This step was taken to ensure that if there are any prosodic cues associated with the different scope readings, they would be provided to the child subjects by the experimenter holding the puppet" (pg. 130). A similar statement is echoed in Lidz & Musolino (2005) (pg. 87-88). It is far from clear, however, what the "most naturally compatible" prosody was in those studies – or, given the results of our production study, what this would be, even from puppeteer to puppeteer. On a related note, Gualmini *et al.* (2008) say in footnote 14 of a

previous study of Gualmini's that the children's lower acceptance rate in one condition "might be due to the particular intonation used by the speaker who recorded the stimuli for that experiment" but say nothing about what this intonational difference might have been, or how it could have produced a difference.

Gualmini *et al.* (2008) do state explicitly that their test sentences, such as (47), in a context supporting the negation> *every* reading were "uttered with the intonation that is required by the inverse scope interpretation in adult English: stress on *every*, de-stressed *wasn't*, and rising intonation on *delivered*" (pg. 219).

(47) Every letter wasn't delivered.

Why they describe the "required" intonation this way is unclear, given (a) the claims in the theoretical literature reviewed in §2, which do not describe the required or favored prosody in this way, and (b) the lack of experimental evidence to this effect at the time that their paper was written. This point aside, children in this condition did access the negation>*every* reading 80% of the time. However, the authors do not attribute the increase in accessing this reading to the prosody, and it is not even possible to identify the contribution of prosody in these experiments, given that this factor was not controlled for by introducing a similar condition in which this prosody was not used, or by ensuring that a consistently falling contour was used for the *every*>negation reading.

Lidz & Musolino (2005) and Musolino & Lidz (2006) refer to the McMahon *et al.* (2004) study in a footnote, as "evidence that adult speakers do not normally use prosody or intonation to indicate the scope of a quantificational subject with respect to negation" (Lidz & Musolino, 2005). However, in §3, we reviewed reasons why we think it is not possible to make this generalization based on that study. Even if it were, though, note that this finding says nothing

about whether the participants in their study may have been sensitive to any prosodic cues that may have been present in the delivery of the sentences. It therefore remains a wide open question whether children are sensitive to the prosodic and acoustic cues associated with quantifiernegation sentences, as the adult participants in our perception studies were.

Building upon a discussion in Musolino & Lidz (2006) (pg. 832), we would like to suggest that the contextual manipulations in the previous studies may have been successful not only in and of themselves, but because they also carried with them prosodic cues that may have made the 'inverse scope' reading more salient (or easier to access in processing). For example, the contrast sentences in that paper (cf. (48)) may have provided children with the necessary contrast in scalar alternatives, coupled with fall-rise prosody needed to access the negation>*every* reading.

(48) Every horse jumped over the log [and/but] every horse didn't jump over the fence. Given that 'log' and 'fence' form members of a scale ('things that the horses could have jumped over'), both should be instantiated with fall-rise prosody – especially given the continuation rise on *log* (cf. Ward & Hirschberg, 1985). Moreover, given the contrast between the conjoined positive and negative statements, negation is surely focused in the second statement, meaning that negation is not part of the presupposition and the QUD is positive (cf. Jackendoff, 1972; Kadmon & Roberts, 1986) – precisely the scenario that favors fall-rise and negation taking wide scope. Of course, the findings of Viau, Lidz, & Musolino (2010) reveal that children can access the inverse scope reading at a comparable rate without the explicit contrast, given other salient contextual manipulations and experiment structuring, so it cannot be the case that the prosodic correlate is *necessary* – an observation consistent with claims by Ward & Hirschberg (1985) reviewed much earlier. A fruitful area for future research, then, would be to try to paint a clearer picture of what the precise role of prosody is when children are able to access the 'inverse scope'/negation>*every* readings that can be so elusive for them. Such research would allow us to see at what stage in development this sensitivity to the auditory correlates of sentence interpretation first manifests itself, and determine how grammatical principles and processing capacity interact in language development. The current research clearly demonstrates that part of becoming an adult language user involves acquiring the ability to recruit prosodic and acoustic cues to access the interpretations of scopally ambiguous sentences in a discourse context.

References

- Akmajian, A., & Jackendoff, R. (1970). Coreferentiality and stress. *Linguistic Inquiry*, *1*, 124-126.
- Baltazani, M. (2002). The prosodic structure of quantificational sentences in Greek. In M.
 Andronis, E. Debenport, A. Pycha, K. Yoshimura (Eds.), *Proceedings from the 38th Meeting of the Chicago Linguistic Society* (pp. 63-78). Chicago: Chicago Linguistic Society.
- Baltazani, M. (2003). *Quantifier scope and the role of intonation in Greek*. Unpublished doctoral dissertation, UCLA.
- Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook, 3*, 15-70.
- Boersma, P., & Weenink, D. (1999-2011). Praat: doing phonetics by computer. [Computer program]. Retrieved from http://www.praat.org/.
- Bolinger, D. L. (1965). Forms of English: Accent, morpheme, order. Ed. by I. Abe & T. Kanekiyo. Cambridge, Mass.: Harvard University Press.
- Büring, D. (1997). The great scope inversion conspiracy. *Linguistics and Philosophy*, 20, 175-194.
- Cedrus Corporation. (2012). SuperLab 4.5 [Stimulus presentation software]. San Pedro, CA.
- Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, Mass.: Harvard University Press.
- Fodor, J.D. (2002a). Prosodic disambiguation in silent reading. In M. Hirotani (Ed.),
 Proceedings of the 32nd Annual Meeting of the North East Linguistic Society (NELS) (pp. 113-132). Amherst, Mass.: GLSA.

Fodor, J. D. (2002B) Psycholinguistics cannot escape prosody. In *Proceedings of the 2002 International Conference on Speech Prosody*. Aix-en-Provence, France, April. Retrieved from http://aune.lpl.univ-aix.fr/~sprosig/sp2002/docs/pdf/fodor.pdf

Frazier, L., & Clifton, C. (1996). Construal. Cambridge, Mass.: MIT Press.

- Gualmini, A., Hulsey, S., Hacquard, V., & Fox, D. (2008). The question-answer requirement for scope assignment. *Natural Language Semantics*, 16, 205-237.
- Hirschberg, J. (1985). A theory of scalar implicature (natural languages, pragmatics, inference).Doctoral dissertation, University of Pennsylvania.
- Hirschberg, J., & Avesani, C. (1997). The role of prosody in disambiguating potentially ambiguous utterances in English and Italian. In A. Botinis (Ed.), *Intonation: Theory, models, and applications* (pp. 189-192). Retrieved from http://www.iscaspeech.org/archive_open/archive_papers/int_97/inta_189.pdf
- Hirschberg, J., & Avesani, C. (2000). Prosodic disambiguation in English and Italian. In A.Botinis (Ed.), *Intonation: Theory, models, and applications* (pp. 87-95). Netherlands: Kluwer Academic Publishers.
- Jackendoff, R. S. (1972). *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
- Jackson, S. (2006, March). Prosodic phrasing and logical scope in English. Paper presented at the 19th Annual CUNY Conference on Human Sentence Processing. New York, NY
- Kadmon, N., & Roberts, C. (1986). Prosody and scope: The role of discourse structure. In A. M.
 Farley, P. T. Farley, & K.-E. McCullough (Eds.), *Proceeding from the 22nd meeting of the Chicago Linguistics Society [Part 2]: Papers from the parasession on pragmatics and grammatical theory* (pp. 16-28). Chicago: Chicago Linguistic Society.

- Koizumi, Y. (2009). *Processing the not-because ambiguity in English: The role of pragmatics and prosody*. Unpublished doctoral dissertation, City University of New York.
- Kitagawa, Y., Fodor, J. (2006). Prosodic influence on syntactic judgments. In Fenselow, Gisbert,
 C. Fery, M. Schlesewsky, & R. Vogel (Eds.), *Gradience in grammar: Generative perspectives* (pp. 336-358). Oxford: Oxford University Press
- Ladd, D. R., Jr. (1980). Intonation and Grammar. *Structure of intonational meaning: Evidence from English*. Bloomington, Indiana: Indiana University Press.
- Lehiste, I., Olive, J. P., & Streeter, L. A. (1976). Role of duration in syntactic disambiguation. *Acoustical Society of America*, 60(5), 1199-1202.
- Liberman, M., & Sag, I. (1974). Prosodic form and discourse function. In M. W. LaGaly, R. A. Fox, and A Bruck (Eds.), *Proceedings of the 10th Regional Meeting of the Chicago Linguistics Society 10* (pp. 416-427). Chicago, IL: Chicago Linguistic Society.
- Lidz, J., & Musolino, J. (2002). Children's command of quantification. Cognition, 84, 113-154.
- Lidz, J. & Musolino, J. (2005). On the status of quantificational indefinites: The view from child language. *Language Acquisition*, *13(2)*, 73-102.
- Martí, L. (2001). Intonation, scope, and restrictions on quantifiers. In K. Megerdoomian and L.A. Barel (Eds.), WCCFL 20 Proceedings (pp. 372-385). Somerville, MA: Cascadilla Press.
- McMahon, E., Lidz, J., & Pierrehumbert, J. (2004). Suprasegmental cues to meaning in childdirected speech. Paper presented at the 17th CUNY Sentence Processing Conference, College Park, MD.
- Musolino, J. (1998). Universal Grammar and the acquisition of semantic knowledge: An experimental investigation into the acquisition of quantifier-negation interaction in

English. Doctoral dissertation, University of Maryland, College Park.

- Musolino, J. (2011). Studying language acquisition through the prism of isomorphism. In J. de
 Villiers and T. Roeper (Eds.), *Handbook of generative approaches to language Acquisition* (pp. 319-349). New York: Springer Science & Business Media.
- Musolino, J., & Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics, 44-4*, 817-852.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Doctoral dissertation, Massachusetts Institute of Technology.
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, 90(6), 2956-2970.
- Roland, D., Yun, H., Koenig, J., & Mauner, G. (2012). Semantic similarity, predictability, and models of sentence processing. *Cognition*, *122*(3), 267-279.
 doi:10.1016/j.cognition.2011.11.011
- Sauerland, U., & Bott, O. (2002). Prosody and scope in German Inverse Linking Constructions.
 In B. Bel and I. Marlien (Eds.), *Proceedings of the Speech Prosody 2002* conference, (pp. 623-628). Aix-en-Provence, France: Laboratoire Parole et Langage.
- Speer, S. R., Crowder, R. G., & Thomas, L. M. (1993). Prosodic structure and sentence recognition. *Journal of Memory and Language*. 32(3), 336-358.
- Syrett, K., Greenberg, L., Kelliher, M., & Kawahara, S. (2011). The acoustic features and perception of children's clear speech. Paper presented at the Annual Meeting of the Linguistic Society of America, Pittsburgh, PA.

Venditti, J., Stone, M., Nanda, P., & Tepper, P. (2002) Discourse constraints on the

interpretation of nuclear-accented pronouns. In *Proceedings of the 2002 International Conference on Speech Prosody*. Aix-en-Provence, France, April.

- Viau, J., Lidz, J., & Musolino, J. (2010). Priming of abstract logical representations in 4-yearolds. *Language Acquisition*, 17, 26-50.
- Ward, G., & Hirschberg, J. (1985). Implicating uncertainty: The pragmatics of fall-rise intonation. *Language*, *61*(4), 747-776.

Appendix A

Items used in Perception Study 1 and 2 are marked by * and indicated as training or test items. Where production and perception comprehension questions differed, both sets are included. Discourse sentences for Perception Study 2 were minimally altered to fit on one line on the computer screen, while maintaining the integrity of the discourse context.

Control Items: Pronominal reference

(1) Alan punched Owen and then he kicked him. *(training, 1 and 2)

a. <u>Context 1</u>: kicker = Alan

Alan is our school's local bully, and picks a new victim every day of the week. Today is Tuesday, which means Owen is in trouble. When the teachers weren't looking, Alan seized the opportunity. Alan punched Owen, and then he kicked him.

b. <u>Context 2</u>: kicker = Owen

Alan is our school's local bully, and he thought he'd pick on the new kid, Owen. Little did Alan know, Owen is a trained kick-boxer. The fight didn't go as Alan expected. When Alan made the first move, Owen struck back. Alan punched Owen, and then he kicked him.

c. <u>Comprehension Question</u>

Production

Who did the kicking?

- a. Alan
- b. Owen

Perception training

Which sentence should follow?

a. Poor Owen! That's really going to hurt tomorrow.

- b. Both boys left bruises on the other.
- (2) Ryan passed Nolan and then he drove off the road. *(training, 2)
- a. <u>Context 1</u>: Ryan drives off the road.

Two boys were street-racing down a narrow road. Ryan was trying to catch up to Nolan, but was paying more attention to the race than to the road. At a bend in the road, Ryan decided to make his move, but he was careless. Ryan passed Nolan and then he drove off the road.

b. <u>Context 2</u>: Nolan drives off the road.

Two boys were street-racing down a narrow road. Ryan was a very aggressive and skilled driver who knew how to take advantage of the situation, but Nolan was new to the game, and couldn't handle sudden moves by other drivers. That explains what happened next. Ryan passed Nolan and then he drove off the road.

c. <u>Comprehension Question</u>

Who drove off the road?

- a. Ryan
- b. Nolan
- (3) Aaron introduced William and then he thanked him.
- a. <u>Context 1</u>: thanker = Aaron

At a recent comic book convention, William was the scheduled keynote speaker. Aaron was the master of ceremonies, in charge of speaker introductions, and was also very appreciative of William coming to give this important talk. Aaron introduced William, and then he thanked him.

b. <u>Context 2</u>: thanker = William

At a recent convention, Aaron had to step in last minute as the master of ceremonies, after the person who was scheduled to fill this role called in sick. William, who was being introduced to give the keynote address, was very appreciative that Aaron—a well-known member of the society—had agreed to take on this role. Aaron introduced William and then he thanked him

c. <u>Comprehension Question</u>

Who did the thanking?

a. Aaron

b. William

(4) Mary admires Arianna but she doesn't like her. *(training, 1)

a. <u>Context 1</u>: Mary doesn't like Arianna

Arianna is a professional singer on Broadway who is known for her attitude as a diva. Mary is an aspiring actress who is an avid fan of Broadway, especially Arianna's work, but she has personal experience with Arianna's nasty attitude backstage. Mary admires Arianna but she doesn't like her.

b. <u>Context 2</u>: Arianna doesn't like Mary

Arianna, a professional singer on Broadway who is known the world over for her ability, takes on even the most challenging of roles with impressive grace and talent. Mary is an up-and-coming actress, who is a big fan of Arianna's work, but at one point she rubbed Arianna the wrong way by mistaking her for another person. Arianna has never forgotten this error, and holds it against Mary to this very day.

Mary admires Arianna but she doesn't like her.
c. <u>Comprehension Question</u>

Which is true?

a. Mary doesn't like Arianna

b. Arianna doesn't like Mary

Control Items: Focus-Sensitive Item Only

- (1) Larry only elbowed Riley. *(test, 1 and 2)
- a. <u>Context 1</u>: focus on VP

Larry and Riley were playing in a soccer match. They both went for the ball, and there appeared to be some contact. Riley went down holding his head in his hands. The crowd was outraged, thinking the injury was more serious than it actually was. But Riley is a good actor, and the camera footage revealed that the injury wasn't that bad. Larry only elbowed Riley. He was glad it wasn't worse.

b. <u>Context 2</u>: focus on DP

In the final seconds of a regulation soccer game, Larry and Riley—bitter rivals from the two opposing teams—were among a group of players crowded around the ball. A header came in their direction, and in the process of getting it, Larry's hand made contact with someone's head. Larry was instantly worried that he had hurt one of his teammates. However, when he looked up and saw who it was, he smiled slyly. Larry only elbowed Riley. He was glad it wasn't someone else

c. <u>Comprehension Question</u>

Production

What sentence would most logically come next?

a. He was glad it wasn't worse.

b. He was glad it wasn't someone else.

Perception

Which sentence should follow?

a. Luckily for Riley, it wasn't worse.

- b. He didn't make contact with anyone else.
- (2) Mary only ran one mile. *(test, 1 and 2)

a. <u>Context 1</u>: focus on VP

Mary and her friends recently participated in a fundraiser race that involved running a mile, biking a mile, and swimming a mile. Mary's friends had participated last year and were regulars at the gym. However, this was Mary's first year, and to be honest, she is not very athletic. After the first course, she couldn't go any farther. Everyone else ran one mile, biked one mile, and swam one mile. Mary only ran one mile. She should have also swam and biked.

b. <u>Context 2</u>: focus on DP (numeral)

This morning, Mary (who used to run every day, but had not done so in years) decided to attempt a 5-mile run. However, a mile into it, she got a serious leg cramp, and had to call her friend to come pick her up. Mary thought she could accomplish the bigger goal of running 5 miles that morning, but her accomplishment was much more modest. Mary only ran one mile. She should have run for longer.

c. <u>Comprehension Question</u>

What sentence would most logically come next?

a. She should have also swam and biked.

b. She should have run for longer.

(3) Warren only likes the Orioles.

a. <u>Context 1</u>: focus on VP

Ben is absolutely crazy about the Orioles baseball team, and has been considering purchasing season tickets. However, he didn't want to go to the games alone, so he asked his friend, Warren, if he is interested in purchasing season tickets, too. Warren declined, since he's not as big a fan of the team as Ben is. Warren only likes the Orioles. He doesn't like them enough to spend the money on season tickets.

b. <u>Context 2</u>: focus on DP

Ben recently won tickets to a Mets-Phillies baseball game from a contest at a local radio station. He wanted to ask one of his buddies to go with him, and was considering asking Warren. Then he realized that it would be absolutely useless to ask Warren to go. Warren only likes the Orioles. He wouldn't be interested in going to any other game.

c. <u>Comprehension Question</u>

What sentence would most logically come next?

a. He doesn't like them enough to spend the money on season tickets.

b. He wouldn't be interested in going to any other game.

Control Items: Focus-Sensitive Item Even

- (1) She even painted the garage. *(test, 1 and 2)
- a. <u>Context 1</u>: focus on VP

Margaret wanted to surprise her husband for his birthday to show how much she appreciated

him. She is not a home improvement person, but decided to transform the garage into a personal workshop for him. She really went all out. She organized it, installed a worktable, added shelves and pegboards, and so on. She even painted the garage. There isn't any more work to be done on the garage.

b. <u>Context 2</u>: focus on DP

Margaret is not really a home improvement sort of person, but recently something has gotten into her. She decided to give the exterior of her home a completely new look. She painted everything—the siding, the shutters, the doors—you name it. She even painted the garage. She is out of things to paint!

c. <u>Comprehension Question</u>

What sentence would most logically come next?

a. She is out of things to paint!

b. There isn't any more work to be done on the garage.

(2) She even composts her newspapers. *(test, 1 and 2)

a. <u>Context 1</u>: focus on VP

Hannah is very environmentally conscious. She gets a lot of newspapers and would never just throw them away after she is done with them. She uses them for papier maché, lines her bird cages with them, lights her outdoor grill with them—you name it. She even composts her newspapers. She is probably thinking of other ways to use newspapers at this very moment.

b. <u>Context 2</u>: focus on DP

Hannah is environmentally conscious. She has a big compost pile in her backyard. Because recycling uses energy, she puts anything that is not reusable into this pile. This includes banana

peels, food scraps and rags. She even composts her newspapers. Everything that can be composted gets composted.

c. <u>Comprehension Question</u>

What sentence would most logically come next?

a. She is probably thinking of other ways to use newspapers at this very moment.

b. Everything that can be composted gets composted.

Control Items: Because clause and Negation

(1) Georgia isn't singing because she is preparing for an audition. *(test, 1 and 2)

a. <u>Context 1</u>: *because* > negation, NOT singing

Georgia and her friends are out at a bar for their weekly karaoke night. Georgia has a beautiful voice, but isn't taking her usual turn at the mike this week. When Simon sees Georgia pass up a chance to sing, he leans over to ask her friends what is going on. They explain the situation to him. Georgia isn't singing because she's preparing for an audition. She has to rest her voice.

b. <u>Context 2</u>: negation > *because*, IS singing

Georgia has a beautiful voice and sings all the time, even in the shower. Her roommate is very familiar with this. When a friend comes over to visit and asks Georgia's roommate if Georgia is practicing for an upcoming role she's trying to land, the roommate explains the situation. Georgia isn't singing because she's preparing for an audition. She just likes to sing.

c. <u>Comprehension Question</u>

Production

Is Georgia singing?

a. Yes

b. No

Perception

Which sentence should follow?

a. She needs to save her voice.

b. She just likes to sing all the time.

(2) They're not late because of his driving. *(test, 1 and 2)

a. <u>Context 1</u>: *because* > negation, NOT late

Mark always drives Andrea and Ben to work. Today they needed to be at the meeting by 9 AM sharp. Unfortunately, there was an accident on the main road, and traffic was backed up for miles. The radio was reporting major delays. But Mark, being extremely clever and a speedy driver, took several shortcuts in order to avoid all the traffic. It's 8:45, and they're already in the conference room, ready for the meeting. They're not late because of his driving. I think they owe Mark a nice lunch.

b. <u>Context 2</u>: negation > *because*, ARE late

Mark always drives Andrea and Ben to work. Mark has always had the reputation of being an overly cautious driver, but lately he's been taking steps to change that, and is actually turning out to be a very aggressive, speedy driver. This morning, Andrea and Ben arrived in the office well after 9 AM, and their boss was not happy. But this time, you can't blame Mark. They're not late because of his driving. Andrea took too much time putting on her make-up.

c. <u>Comprehension Question</u>

Production

Are they late?

a. Yes

b. No

Perception

Which sentence should follow?

a. He's a very fast driver.

b. The car had a flat tire.

(3) Omar isn't in shape because he runs outdoors. *(test, 1 and 2)

a. <u>Context 1</u>: *because* > negation, NOT in shape

Omar usually runs a three-mile circuit in the woods near his house. He's somewhat insistent that the best training takes place outside in the elements, instead of in some fancy, high-priced gym. Lately, however, the weather has been horrendous, and Omar hasn't been able to run at all. His training has really suffered. Omar isn't in shape because he runs outdoors. If he trained at a gym, he wouldn't have a problem.

b. <u>Context 2</u>: negation > *because*, IS in shape

Omar runs the same three-mile circuit in the woods by his house every day, and is very fit. His roommate Ryan wants to get in shape, and mistakenly credits Omar's success to his exercising outside, failing to realize that Omar runs three miles every day. Ryan thinks that if he runs a mile once a week and does it outside, he'll be as buff as Omar. But that's not sound logic. Omar isn't in shape because he runs outdoors. He's in shape because he runs every day.

c. <u>Comprehension Question</u>

Production

Is Omar in shape?

a. Yes

b. No

Perception

Which sentence should follow?

a. It's been too cold and wet lately.

b. He just loves to do kickboxing.

(4) Elaine isn't laughing because she's embarrassed.

a. <u>Context 1</u>: *because* > negation, NOT laughing

Elaine giggles at just about anything, but not when she's put on the spot. Mario, who tends to tell off-color jokes, came over for dinner tonight and made a joke at her expense. Just as I expected, his jokes have made her feel uncomfortable. Elaine isn't laughing because she's embarrassed. Now she's just sitting there.

b. <u>Context 2</u>: negation > *because*, IS laughing

At office parties, Mario likes to tell off-color jokes and awkward stories about his co-workers. Elaine has always thought that Mario would get in trouble for this behavior, and told him she would revel in it when he finally crossed the line. This evening, Mario went too far and offended his boss, and tried to recover somewhat sheepishly. Elaine thought this was hilarious. Elaine isn't laughing because she's embarrassed. She knew this day would come.

c. <u>Comprehension Question</u>

Is Elaine laughing?

a. Yes

b. No

(5) Ryan didn't win the trophy because he took steroids.

a. <u>Context 1</u>: *because* > negation, DID NOT win

Ryan has been training hard for months in preparation for a big marathon. His friend suggested he use steroids to get an edge on the competition. Ryan hesitated, but finally decided to do it, thinking he wouldn't get caught. He ended up placing first in the race but then a drug test disqualified him. Ryan didn't win the trophy because he took steroids. It was the worst mistake of his life.

b. <u>Context 2</u>: negation > *because*, DID win

Ryan is training hard for a big marathon. His friend suggested he try steroids to get an edge on the competition. Ryan thought about this for a long time, and although this would have ensured a win, Ryan trained naturally instead, and still came in first place. Ryan didn't win the trophy because he took steroids. All of his training just paid off in the end.

c. <u>Comprehension Question</u>

Did Ryan win the trophy?

a. Yes

b. No

Test Sentences: Universal Quantifier All and Negation

- (1) All the magnolias won't bloom. *(test; contexts 1 and 2, perception 1 and 2)
- a. <u>Context 1</u>: *all* > negation, negation associated with presupposition

The township decided to plant magnolia saplings a number of years ago to line a path through the park. They have experienced lovely blossoms every year. However, this year the area is experiencing less-than-standard rainfall, which means that they expect the magnolias to struggle this year, with only a few surviving. In fact, I think the situation is much more dire than that. All the magnolias won't bloom. They'll just have to wait till next year.

b. <u>Context 2</u>: negation > *all*, negation associated with focus/assertion

A few years ago, the township decided to plant magnolia saplings to line a path through the park. The saplings on the north side were planted mainly in sand and haven't been getting nearly enough nutrients. However, the soil near the south side is rich, and the magnolias are thriving there. All the magnolias won't bloom. But I bet the ones on the south side will.

c. <u>Context 3</u>: *all* > negation , negation associated with presupposition, scalar contrast with DP (magnolias)

An aggressive beetle that targets magnolia trees is spreading through our area, and the magnolias are doomed. The township has been planning on taking pictures for their website next month. The official photographer is concerned that there won't be beautiful rows of trees in the background for his pictures. I think he's worrying too much. All the magnolias won't bloom. However, there will still be other trees that will look just as lovely.

d. <u>Context 4</u>: negation > *all*, negation associated with focus/assertion, scalar contrast with quantity

The weather recently has been conducive to plant growth, and all the trees are looking healthy. Some optimistic members of the township are predicting that each of the magnolia trees will give us lovely, fragrant blossoms to enjoy all season. But I think they're being rather unrealistic, and I keep telling them this. All of the magnolias won't bloom. The odds of each of them blooming are pretty slim.

e. <u>Comprehension Question</u>

Production

Will any magnolias bloom?

a. Yes

b. No

Perception

Which sentence should follow?

a. The situation is dire for the newly-planted saplings.

b. But I bet the ones on the south side will.

(2) All the wool lining wasn't worn. *(test; contexts 1 and 2, perception 1 and 2)

a. <u>Context 1</u>: *all* > negation, negation associated with presupposition

Mandy was in need of a heavy winter jacket but had a limited budget. She was hoping to find one when she went to the thrift store, even though she knew there would be a chance that some of the lining would be in need of repair. Eventually, she found a nice, warm jacket. When she looked inside, she couldn't believe how lucky she was. All the wool lining wasn't worn. The mission was a huge success!

b. <u>Context 2</u>: negation > all, negation associated with focus/assertion

With the weather turning colder, Mandy was going through her closet looking for her winter coat. She thought she had remembered that the lining on this particular jacket was in pretty bad condition, and it would all have to be removed. When she found it, she was pleasantly surprised. All the wool lining wasn't worn. Only the sleeves needed repair.

<u>Context 3</u>: *all* > negation , negation associated with presupposition, scalar contrast with
DP (adjective wool)

Mandy's roommate wanted to go shopping for a new coat at a thrift store. Mandy went ahead to scout it out. She found coats with various types of linings. Each one of the coats with wool lining was in excellent condition, but the ones with the polyester lining were less pristine, to say the least. When she got home, Mandy gave her roommate the update. All the wool lining wasn't worn. However, the polyester lining was in tatters.

d. <u>Context 4</u>: negation > *all*, negation associated with focus/assertion, scalar contrast with quantity

Mandy was shopping for a winter coat at a thrift store. Her roommate warned her that she would find that the wool lining in the coats would be in absolutely terrible condition. Even though Mary found some coats that fit this description, others actually had pristine linings. When she came home, her roommate said, "I told you so! They were all worn, weren't they?" Mary responded calmly. All the wool lining wasn't worn. Wait until you see the fantastic coat I found.

e. <u>Comprehension Question</u>

Production

Was any of the wool lining not worn?

a. Yes

b. No

Perception

Which sentence should follow?

a. The coat was in perfect condition.

- b. Only the sleeves needed repair.
- (3) All the moms didn't allow eyeliner. *(test; contexts 1 and 2, perception 1 and 2)

a. <u>Context 1</u>: *all* > negation, negation associated with presupposition

Some of the girls in the neighborhood decided to throw a party, where they would help each other apply makeup in preparation for the upcoming dance. The girls anticipated that some of their moms wouldn't let them wear eyeliner. It turns out that the moms were all on the same page. All the moms didn't allow eyeliner. This didn't come as a real surprise.

b. <u>Context 2</u>: negation > all, negation associated with focus/assertion

Several moms were helping their daughters get ready for the upcoming school dance. This is a progressive school, and moms are usually lenient about certain things, so even the younger girls thought their moms would approve of eyeliner. But at the dance only the older girls were wearing it. All the moms didn't allow eyeliner. Only the moms of the older girls let their daughters wear it.

<u>Context 3</u>: *all* > negation , negation associated with presupposition, scalar contrast with DP (eyeliner)

Some of the girls in the neighborhood decided to throw a party, where they would help each other apply makeup in preparation for the upcoming dance. The girls anticipated that some of their moms would express concern about the girls wearing eye makeup. The situation was more nuanced than the girls expected. All the moms didn't allow eyeliner. But they were fine with them wearing mascara.

d. <u>Context 4</u>: negation > *all*, negation associated with focus/assertion, scalar contrast with quantity

Recently, several moms helped their 5th grade daughters get ready for a school dance. Girls being girls, they all wanted to wear make-up. The moms were generally fine with them wearing a little blush—maybe even some light eyeshadow or some lip gloss. But some moms drew the line there. All the moms didn't allow eyeliner. Those girls will have to wait until they're in middle school for that.

e <u>Comprehension Question</u>

Production

Did any moms allow eyeliner?

a. Yes

b. No

Perception

Which sentence should follow?

- a. They were all in agreement.
- b. Only the moms of the older girls let their daughters wear it.

(4) All the newlyweds didn't buy ironware.

a. <u>Context 1</u>: *all* > negation, negation associated with presupposition

Past sales at Macy's have shown that a small percentage of newlyweds usually purchase ironware. However, times are changing, and ironware is not as popular as it used to be. In fact, no one's really interested in it anymore. Indeed, when an employee checked the store records for this past year, he confirmed this pattern. All the newlyweds didn't buy ironware. They opted for kitchen appliances instead.

b. <u>Context 2</u>: negation > all, negation associated with focus/assertion

Past sales at Macy's have shown that ironware is an extremely popular item for newlyweds; in fact, it's very rare to find newlyweds who don't purchase it! So employees were certain that the store sales for this past year would reflect this trend. However, when they reviewed the records,

they discovered that this wasn't the case. All the newlyweds didn't buy ironware. Only a small fraction did.

<u>Context 3</u>: *all* > negation , negation associated with presupposition, scalar contrast with *DP (ironware)*

Two employees at Macy's placed a bet to see which item each set of newlyweds would buy. Gus said they would all purchase ironware, but Sam insisted they would all purchase kitchen appliances. It turns out Sam was right. All the newlyweds didn't buy ironware. They bought appliances...and now Gus owes Sam \$50.

d. <u>Context 4</u>: negation > *all*, negation associated with focus/assertion, scalar contrast with quantity

I recently overheard a Macy's employee discussing what newlyweds purchased with all the cash they got after their weddings. He had noticed a growing trend in their purchases, and said that they all bought things from the kitchen department. I had expected to hear him say that ironware was the most popular item. But that's not what he said. All the newlyweds didn't buy ironware. There were too many other things to choose from.

e. <u>Comprehension Question</u>

Did any of the newlyweds buy ironware?

- a. Yes
- b. No

Test Sentences: *Many/Most* and Negation

- (1) Liam doesn't know many alumni. *(test; contexts 1 and 2, perception 1 and 2)
- a. <u>Context 1</u>: *many* > negation, negation associated with presupposition

The alumni association is looking for a new president who is going to be able raise money. Todd nominated Liam. However, I think that's a bad idea. Liam doesn't know many alumni. He won't be able to bring in a lot of money.

b. <u>Context 2</u>: negation > *many*, negation associated with focus/assertion

The alumni association is looking for a new president who is going to be able raise money. Todd nominated Liam. I think that is a great idea. Liam doesn't know many alumni. But the ones he does know have deep pockets.

c. <u>Context 3</u>: negation > *many*, embedded clause favoring falling contour

The alumni association is looking for a new president who is going to be able to raise money. Todd nominated Liam. I seconded this nomination, saying that Liam knows a handful of wealthy alumni. However, Carl voiced his concern that Liam doesn't have a lot of connections. Carl's right that Liam doesn't know many alumni. However, the few rich ones he knows will bring in the big bucks.

d. <u>Comprehension Question</u>

Production

Which sentence would more logically come next?

a. He won't be able to bring in a lot of money.

b. But the ones he does know have deep pockets.

Perception

Which sentence should follow?

- a. He really has to make more connections.
- b. But the ones he knows are well established in the community.

(2) Neil doesn't enjoy most musicals. *(test; contexts 1 and 2, perception 1 and 2)

a. <u>Context 1</u>: *most* > negation, negation associated with presupposition

Neil is an avid fan of the theater. I bought tickets to the musical "Chicago" for him as a gift. My friend Adam was concerned about the choice. I realized he was right. Neil doesn't enjoy most musicals. He thinks they are very cheesy.

b. <u>Context 2</u>: negation > *most*, negation associated with focus/assertion

Neil is an avid fan of the theater. I bought tickets to the musical "Chicago" for him as a gift. My friend Adam was concerned about the choice. I had to assure him that it was ok. Neil doesn't enjoy most musicals. But I know for a fact that he adores "Chicago".

c. <u>Context 3</u>: negation > *many*, embedded clause favoring falling contour

Neil is an avid fan of Broadway. I bought tickets to the musical "Chicago" for him as a gift. My friend Adam was concerned about the choice. He said that Neil doesn't like going to musicals. However, I know for a fact that there are certain ones he does appreciate. I acknowledged that Neil doesn't enjoy most musicals. But I know for a fact that he adores Chicago.

d. <u>Comprehension Question</u>

Which sentence would more logically come next?

a. He thinks they are very cheesy.

b. But I know for a fact that he adores Chicago.

(3) Melanie doesn't follow most rules.

a. <u>Context 1</u>: *most* > negation, negation associated with presupposition

At the beginning of the year, my roommates and I made a list of ground rules for the apartment.

We are having a meeting tonight to discuss Melanie. She has been a sore spot for all of us.

Melanie doesn't follow most rules. We need to ask her to leave.

b. <u>Context 2</u>: negation > *most*, negation associated with focus/assertion

At the beginning of the year, my roommates and I made a list of ground rules for the apartment. We are having a meeting tonight to discuss Melanie. She is having some difficulty following along, but she means well. Melanie doesn't follow most rules. But at least she remembers to take out the trash.

c. <u>Context 3</u>: negation > *many*, embedded clause favoring falling contour

At the beginning of the year, my housemates and I made a list of ground rules for the apartment. Melanie has been following some rules, but she's lax about the others. One of my housemates said that we should kick Melanie out, since she blatantly disregards all the rules. I said that it's true that Melanie doesn't follow most rules. But at least she remembers to take out the trash.

d. <u>Comprehension Question</u>

Which sentence would more logically come next?

a. We need to ask her to leave.

b. But at least she remembers to take out the trash.

(4) Hermione doesn't believe in many omens.

a. <u>Context 1</u>: *many* > negation, negation associated with presupposition

Hermione is a very practical person and doesn't bother much with superstitions. Someone tried to scare her by shooing a black cat in her path. It didn't work. Hermione doesn't believe in many omens. They'll have to find some other way to scare her.

b. <u>Context 2</u>: negation > *many*, negation associated with focus/assertion

Hermione is a very practical person and doesn't bother much with superstitions. Someone tried

to scare her by shooing a black cat in her path. I thought that could actually work. Hermione doesn't believe in many omens. But she was definitely scared by this one.

c. <u>Context 3</u>: negation > *many*, embedded clause favoring falling contour

Hermione is a very practical person and doesn't bother much with superstitions. I told my friend about her, but he still wanted to try to scare her. He assumed she would still believe in a few and tested this by shooing a black cat in her path. I had said that Hermione doesn't believe in many omens. But she was definitely scared by this one.

d. <u>Comprehension Question</u>

Which sentence would more logically come next?

- a. They'll have to find some other way to scare her.
- b. But she was definitely scared by this one.

Appendix B

Pitch tracks representing minimal pair members for target sentences presented to hearers in

Perception Experiments 1 and 2.







(6) Target sentence: They're not late because of his driving.