J. Linguistics **50** (2014), 453–493. © Cambridge University Press 2014
doi:10.1017/S0022226714000012 First published online 14 February 2014

Prosodic disambiguation of scopally ambiguous quantificational sentences in a discourse context[1]

KRISTEN SYRETT

Rutgers, The State University of New Jersey – New Brunswick

GEORGIA SIMON

Rutgers, The State University of New Jersey – New Brunswick

KIRSTEN NISULA

The Ohio State University

(Received 30 July 2012; revised 17 July 2013)

Researchers have long sought to determine the strength of the relation between prosody and the interpretation of scopally ambiguous sentences in English involving quantification and negation (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), in subsequent work, researchers (e.g. Ladd 1980; Ward & Hirschberg 1985; Kadmon & Roberts 1986) disentangled the link between prosody and scope. Even though these pragmatic accounts predict variability in production, they still allow for some correlation between scope and prosody. To date, we lack systematic evidence to bear on this discussion. Here, we present findings from two perception experiments aimed at investigating whether prosodic information – including, but not limited to, sentence-final contour – can successfully disambiguate such sentences. We show that when speakers provide consistent auditory cues to sentential interpretation, hearers can successfully recruit these cues to arrive at the correct interpretation as intended by the speaker. 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 – quantificational and otherwise – should carefully control for prosody.

1. Introduction

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

(1) **All** the men did**n't** go.

[1] We gratefully acknowledge a Rutgers University startup grant to K. Syrett and a research grant from the Aresty Research Center at Rutgers University. This work benefitted from technical support from the Rutgers Phonology and Field Research Laboratory, discussions with Shigeto Kawahara, contributions from Stephen Klimashousky, and helpful feedback from three anonymous *Journal of Linguistics* reviewers. We thank audiences at ETAP 2 (2011), NELS 43 (2012), and The University of Pennsylvania for their helpful comments and observations.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Under one interpretation – the one captured by the formal logic in (2a) below – 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 ∀) takes scope over VP-level negation (marked by ¬). The second interpretation – the one in (2b), where negation (¬) takes wide scope over the universal quantifier (∀) – indicates that it is NOT the case that all of the men went. The possibility is left open that only some (but not all) of the men went, or that none went – although the 'some' interperpretation is the more salient of the two.

(2) (a) $\forall x.\ man(x) \rightarrow \neg go$
(b) $\neg\forall x.\ man(x) \rightarrow go$

Building on work by Bolinger (1965), Jackendoff (1972) proposed that speakers use prosody to disambiguate sentences such as the one above (Jackendoff's (8.159)–(8.162) collapsed). Specifically, he claimed that the interpretation in (2a) (where negation has narrow scope under the quantifier, and the scopal relation is signaled by '>' for 'takes scope over') is indicated with a sentence-final falling contour (Bolinger's Accent A), as in (3), while the interpretation in (2b) (where negation takes wide scope) has a rising or FALL-RISE contour, as in (4) (which Jackendoff referred to as Bolinger's 'B accent'). For Jackendoff, this difference was encoded in the representation of these sentences, and should therefore be minimally variable.

(3) ∀>¬ ('none')
ALL the men didn't go.
('A accent')

(4) ¬>∀ ('not all')
ALL the men didn't go.
('B accent')

Indeed, it has been claimed that prosody is intimately tied to scope and therefore to sentence interpretation in languages such as German (see Jacobs 1984; Büring 1997, 2003; Krifka 1998; Martí 2001; Sauerland & Bott 2002; Schäfer 2004; Braun 2006). For example, the sentence in (5) from Büring (1997: 177, 180–184), uttered with the fall-rise prosody (as indicated by him with / \ notation), unambiguously favors negation taking wide scope.

(5) **Alle** Politiker sind **nicht** korrupt.
all politicians are not corrupt
'All politicians are not corrupt.'
/ALLE Politiker sind NICHT\ korrupt

A somewhat similar pattern also appears to hold in Greek (see Baltazani 2002, 2003).

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

In English, there is by now considerable evidence that prosody plays a role in sentence interpretation elsewhere (see Fodor 2002a, b). Experimental findings cover a range of lexical items and interpretations including, but not limited to, structural ambiguity (Lehiste 1973; Lehiste, Olive & Streeter 1976; Beach 1991; Price et al. 1991; Speer, Crowder & Thomas 1993; Nagel et al. 1996, Kitagawa & Fodor 2006), pronominal reference (Akmajian & Jackendoff 1970; Hirschberg & Avesani 1997, 2000; McMahon, Pierrehumbert & Lidz 2004), parenthetical vs. integrated content (Price et al. 1991), filler–gap processing (Nagel, Shapiro & Nawy 1994), the scope of focus-sensitive operators *only* and *even* (Hirschberg & Avesani 1997, 2000), contrast in English (and Dutch) constituency (Braun & Chen 2010), and the interaction of a *because* clause with negation (Cooper & Paccia-Cooper 1980; Hirschberg & Avesani 1997, 2000; Koizumi 2009). However, a conspicuous gap still exists when it comes to addressing Jackendoff's seminal claims for quantificational sentences with negation in English.

Filling this gap is important not only for the benefit of linguistic theory and psycholinguistic methodology, but also because it is worthwhile from the perspective of language acquisition to discover the range of surface-level information available in the speech stream that is linked to sentence meaning, and further to determine whether such information is used in sentence disambiguation. To the extent that experimental participants find these auditory cues informative about the speaker's intended interpretation, we have identified a characteristic of what it means to demonstrate adult-like sentence processing. Moreover, if these auditory cues can indeed be perceived as favoring one interpretation over another, and if there is a correlation between certain sentential interpretations and prosody, then this information should be systematically controlled for to put children in the best possible position for success in experimental tasks involving sentential ambiguity.

Previous studies on this particular phenomenon (McMahon et al. 2004; Jackson 2006) 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 reasons. First, Jackendoff's emphasis was on how different instantiations of a single sentence varied with respect to their sentence-final contour (e.g. (3)–(4) above). But as Jackson's (2006) data suggest, prosodic and acoustic differences may reside in other auditory cues, such as word duration, and are also highly variable within and across speakers (see also Syrett, Simon & Nisula, to appear). Second, a number of researchers since Jackendoff (1972) have moved the phenomenon into the purview of pragmatics, arguing that prosody highlights information in the discourse context, which in turn is informative about scopal relations (see Ladd 1980, 1996/2008; Ward & Hirschberg 1985; Kadmon & Roberts 1986; Baltazani 2002, 2003). Studies that do not take into account aspects of the discourse context such as the presence and type of scalar alternatives (Hirschberg 1985), the role of negation in the presupposition or assertion, and the Question Under Discussion (Roberts 1996) (see next section) may not be in a position to uncover the relevant patterns. Moreover, these patterns may also vary due to other features of the context.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

In this paper, we briefly present findings from a Production experiment (Syrett et al. to appear) as a backdrop to reporting the findings of two new Perception experiments aimed at investigating whether participants recruit auditory cues in the service of disambiguating quantificational sentences with negation. To our knowledge, this is the first systematic investigation of this phenomenon in English. In Section 2, we present theoretical claims concerning the role of intonation in the disambiguation of sentences with negation and quantification, beginning with Jackendoff (1972). In Section 3, we review the previous experimental work bearing on this issue. In Section 4, we present the methodology and key findings from our Production experiment, from which the items used in our two Perception experiments (presented in Sections 5–6) were drawn. Previewing our results, we demonstrate that hearers can use auditory cues provided by the speakers to retrieve the intended discourse-supported interpretation. Finally, in Section 7, we conclude with a summary of our main contributions, connecting the findings to ongoing research in the field of language acquisition.

2. Theoretical background

Jackendoff (1972: 352) said of the sentence in (6) that it as a 'a contrast in meaning … produced by a difference in the choice of pitch accent [read, intonation]'.

(6) **All** the men did**n't** go.

When the reading in (7) is intended, negation is associated with the presupposition, giving rise to the representation in (8). That is, it was expected that some quantity of men did not go, and what is asserted is that that quantity is ALL of the men.

(7) ∀>¬ ('none')
ALL the men didn't go.
('A accent')

(8) A accent (fall): negation is part of the **presupposition**, not the focus
Presupposition: λ*Q* [*Q* of the men did**n't go**] is well-formed/under discussion
Assertion: all ∈ λ*Q* [*Q* of the men didn't go]

When the reading in (9) is intended, however, negation is associated with the focused quantifier (*all*), giving rise to the presupposition and assertion in (10). That is, there is some expectation that SOME men went, and what is asserted is that that number is NOT ALL.

(9) ¬>∀ ('not all')
ALL the men didn't go.
('B accent')

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(10) B accent (rise): negation is associated with the **focus** (i.e., the **assertion**), not presupposition
Presupposition: λQ [Q of the men **went**] is well-formed/under discussion
Assertion: all $\notin \lambda Q$ [Q 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 (seen in (11)), *because* (see the unfortunate example in (12)), and focused constituents elsewhere.

(11) I didn't see ALL of the men.
(Jackendoff (1972: exx. (8.181)–(8.182))

(12) Max doesn't beat his wife because he LOVES her.
(Jackendoff (1972: exx. (8.185)–(8.156))

For example, Jackendoff claimed that in both versions of the ambiguous sentence in (13) (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. (Jackendoff 1972: 259–260)

(13) (a) FRED doesn't write poetry in the garden.
'It is Fred who doesn't write poetry in the garden.'
(b) FRED doesn't write poetry in the garden.
'It isn't Fred who writes poetry in the garden.'

Jackendoff (1972: 260) further argued that 'the difference between them always appears at the end of the sentence'.

Liberman & Sag (1974) were among the first to attempt a revision of Jackendoff's account, arguing instead that with sentences 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 (14), accompanied by a fall-rise contour, may be thought of as a rejection addressed to someone who has claimed that elephantiasis is incurable.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(14) Elephantiasis isn't incurable.

While Ladd (1980) agreed with the general direction of Liberman & Sag's (1974) 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 noticeably different. For example, in response to someone who thinks that elephantiasis is fatal, a speaker might reply with (14), with a high pitch accent toward the beginning of *EL* (the contradiction contour). However, in response to someone concerned about his imminent death as a result of contracting elephantiasis or rabies, a speaker might reply with (14) with a high pitch accent on *TI* (the fall-rise/contrastive contour). Ladd also pointed out that while the contradiction contour CANNOT be embedded (as Liberman & Sag had observed), a fall-rise contour CAN be. The implication, then, is that sentences such as *All the men didn't go* could result in a rising pattern for different reasons.

Ladd (1980) thus shared the 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 the fall-rise contour is 'focus within a given set'. An example dialogue from Ladd (1980: 153 ex. (16)), presented here as (15), 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$^{\check{}}$ notation) indicates that s/he fed the cat, and cats are members of the set evoked by A.

(15) A: Did you feed the animals?
B: I fed the$^{\check{}}$ cat.

That the sets are contextually relevant and can be ad-hoc is illustrated by other examples, such as (16) (Ladd's (1980) example (24)).

(16) A: What would you think of getting [–] a dog?
B: A$^{\check{}}$ stove, maybe.

As Ladd (1980: 154) pointed 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 that the speaker and hearer could perhaps agree upon. (See his example (25), page 154.) Ladd's reasoning behind the *all* example is as follows: 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, a scopal relation where negation takes scope over the universal quantifier will be associated with a fall-rise (and not a falling) contour.

Ward & Hirschberg (1985) presented what is easily the most extensive pragmatic account of fall-rise contour to date. They first argued (pages 754–755) that Ladd's (1980) conception of 'focus within a given set' cannot be correct because of examples illustrating that B does not have to evoke membership WITHIN the set mentioned by A. For example, in (17) (their (19)), A is asking about the

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

route that B took traveling in Philadelphia. The dialogue is felicitous, they argued, if B does not know that Walnut ends at 34th. What is not at issue is membership in the set of streets. (The \ / notation was used by Ward & Hirschberg to indicate the syllable hosting the fall-rise accent pattern.)

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

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

(18) 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 (19) (their (17)).

(19) 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 (15) above is still interpretable with falling intonation on *cat*.

Ward & Hirschberg (1985) argued that 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 (see Hirschberg 1985), and uses this contour, because s/he is uncertain about (i) whether it is appropriate to evoke a scale at all, (ii) if some scale is appropriate, which scale should be chosen, or (iii) given a scale, which value (or scalar alternative) should be chosen.

Ward & Hirschberg (1985) noted that because in their account the contour is indicative of a speaker's knowledge of contextually-relevant scales and scalar alternatives, a fall-rise 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 quantificational sentences in question. First, it should be possible to observe a fall-rise contour WITHOUT negation taking wide scope over the quantifier (that is, with *all* taking wide scope over negation). Ward & Hirschberg provided the example in (20) below, in which B responds to A's utterance using a fall-rise contour. Here, 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.

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

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Second, we predict that we should be able to observe negation taking wide scope over the quantifier WITHOUT a fall-rise contour (i.e., with a falling contour). As an example, Ward & Hirschberg pointed out that the embedded clause in (21) (following *I know that*) can be uttered with a falling contour while still favoring negation taking wide scope.

(21) 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 interpreation and an appeal to a pragmatic account is echoed in Kadmon & Roberts (1986), whose main example is presented in (22), with the prosody–interpretation correlations and scopal relations in (23).

(22) He does**n't** hate **most** of the songs.
(23) (a) falling contour in response to 'negative' question
scope: *most* > negation
(b) fall-rise contour in response to 'positive' question
scope: negation > *most*

We note in passing that it is possible to find similar attested examples in an online search:

(24) Although they did**n't** win **many** 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.
('Led Zeppelin and the Recording Connection Audio School', http://www.recordingconnection.com/artists/led-zeppelin)
(25) AMC's superb "Mad Men" did**n't** win **many** awards on the night, but the two they did counted . . .
('Sepinwall on TV: Recapping the Emmys', http://www.nj.com/entertainment/tv/index.ssf/2008/09/sepinwall_on_tv_recapping_the.html)

Kadmon & Roberts (1986: 16) acknowledge that the prosody of a given sentence does indeed appear to favor one interpretation over another, but that these different prosodic patterns also differ with respect to the contexts in which they occur, and the explicit or implicit topical question they address – the QUESTION UNDER DISCUSSION (QUD) (using terminology developed by Roberts 1996). In a nutshell, their claim is that prosody does not directly determine scopal relations. It conveys partial information about the structure of the discourse, and this structure determines scope. The hearer is then left to reconstruct

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

the best fitting (and simplest) interpretation, and 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 be more likely to favor negation taking wide scope over the subject quantifier, such a contour is neither necessary nor sufficient for this interpretation. A certain degree of variability should also be expected, if prosody is tied to speaker uncertainty in a discourse context and to implicit questions that are part of the information structure. Thus, while there may be a correlation between sound and meaning with these ambiguous sentences, there is not a one-to-one mapping, especially between sentence-final contour and scopal relation. One wonders whether – in the face of this variability – hearers can still associate certain prosodic information with certain interpretations (*qua* scopal relations). In the next section, we review relevant experimental work in this area, which suggests that hearers are attuned to auditory correlates to sentence interpretation and can use prosody to reconcile ambiguity in other cases. These findings suggest that a similar pattern of results could be identified for our target sentences involving quantification and negation.

3. Experimental background

A number of studies over the years (covering a range of phenomena and employing a variety of paradigms) have investigated the role of prosody in sentence disambiguation. As we will show, the general picture that emerges is that experimental participants do assign some weight to prosody in determining the interpretation of a potentially ambiguous string, but how much weight they assign varies among speakers, hearers, and contexts. And importantly, a conspicuous gap exists when it comes to systematically probing the quantificational sentences that originally interested Jackendoff and others.

In a set of three studies, Speer et al. (1993) showed that participants use prosody to disambiguate structurally ambiguous sentences and to determine the focus or presuppositional structure. In one study, for example, 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 #* vs. *Either Sam or Susan # and Lara #*) and differences in syntactic structure (e.g. *They are cooking APPLES* vs. *They are COOKING apples*). Participants were sensitive to the prosodic pattern when choosing their paraphrase, but more so for syntactic ambiguities than for placement of focus, as indicated by a calculation

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

of difference scores using hits and false alarms based on participants' choice of intended meaning.

Price et al. (1991) recorded professional radio announcers reading a range of structurally ambiguous sentences that were embedded in discourse contexts. 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. They found considerable variation among the speakers, as well as variability in participants' success rate based on the speaker and the sentence types, with the overall percentage correct reported ranging from 54% to 96%. Listeners were largely successful in distinguishing between apposition and NP attachment (*The neighbors who usually read* [*the dailies* / , *the Daleys*,] *were amused*), with a 91–92% correct rate reported. However, their performance with attachment ambiguities (e.g. *Raoul murdered the man* [*with a gun*]) was less consistent, with the percentage correct ranging between 63% and 78%, and clausal ambiguity (e.g. *Mary was amazed and Dewey was angry*/*Mary was amazed Ann Dewey was angry*), with the percentage correct ranging from 54% to 88%. These rates are important to keep in mind when evaluating the results of our Perception experiments later in the paper.

Other studies have focused more specifically on scopal ambiguity with negation. For example, Cooper & Paccia-Cooper (1980) presented participants with ambiguous target sentences such as the 'negation–*because*' sentence in (26), followed by a disambiguating follow-up sentence or with the target sentence embedded in a disambiguating paragraph.

(26) Dick did**n't** fly the kite **because** it was a beautiful day.

Participants significantly lengthened the vowel of the word preceding the *because* clause (here, *kite*) when producing the interpretation where *because* took wide scope over negation (i.e., 'It was because of the day that Dick didn't fly the kite'). The length of the vowel was approximately 30 ms longer relative to the length of the vowel when negation took wide scope over the *because* clause. This effect, however, only surfaced when participants were made aware of the ambiguity with isolated target sentences.

Hirschberg & Avesani (1997, 2000) presented English and Italian speakers with a range of ambiguous sentences, including negation interacting with a *because* clause (e.g. *William isn't drinking because he's unhappy*), attachment ambiguities, focus-sensitive operators (*only, even*), and the scope of *none*. 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 the *none* sentences or for attachment ambiguities. They did, however, produce differences for the 'negation–*because*' sentences, as captured in Table 1.

The contour data in the far right column can also be re-cast by saying that of the 13 sentences in which there was no falling contour (adding the differences

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Scopal relation	Internal phrase boundary	Falling contour
because > negation	12/18	15/18
negation > *because*	2/18	8/18

Table 1
Pattern of results for 'negation–*because*' sentences in Hirschberg & Avesani (1997, 2000).

of 3 and 10), 10 were observed with negation taking wide scope over *because*. Thus, a non-falling contour was more likely when negation took wide scope. For *only*/*even* sentences, participants regularly placed the nuclear stress in the focused phrase in the sentence.

Further evidence for the role of prosody in disambiguating 'negation–*because*' sentences in English comes from Koizumi's (2009) self-paced silent reading studies. In one experiment, participants were shown sentences that favored one of two possible interpretations, as in (27), then answered a question about the sentence.

(27) (a) Jane did**n't** purchase the white blouse **because** it had a stain.
(*because* > negation)

(b) Jane did**n'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, which neutralized the intonational difference between the two sentences. When a line break was inserted before the *because* clause (which had the effect of inserting a prosodic break before *because*), participants displayed a preference for the interpretation where *because* took wide scope.

With the above studies as background to how experimental participants interpret the role of prosody in sentence ambiguation, we now turn to ambiguous sentences involving quantification and negation, our target sentences in this study. Baltazani's (2002, 2003) complementary production and perception studies in Greek offer the first strong suggestion that speakers and hearers use prosody to disambiguate such sentences. Baltazani (2002) focused on three types of items, captured in English in (28).

(28) (a) He's **not** watching TV **because** he's bored. (negation–*because*)

(b) They did **not** eat **many** apples. (negation and object quantifier)

(c) **Three** nurses helped **every** doctor. (subject and object quantifiers)

The tokens for each item type varied SVO/SOV word order, since Greek allows for variable word order. In the production study, participants viewed a question

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

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 (see (29)).

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

These productions were then presented to participants in a forced choice task. Participants heard the excised target sentence in isolation and were asked to choose a suitable answer.

Baltazani (2002) found that nearly every one of her Greek speakers reliably produced ambiguous sentences with negation with two distinct contours related to the two interpretations. Moreover, hearers were by and large able to correctly identify their corresponding interpretations. When presented with sentences with an OV word order in which the quantifier was focused, hearers were more likely to interpret the quantifier taking wide scope over negation (51% quantifier > negation vs. 31% negation > quantifier). By contrast, when presented with VO sentences in which negation was focused, hearers were more likely to interpret negation taking wide scope over the quantifier (62% negation > quantifier vs. 29% quantifier > negation). Interestingly, however, while sentences involving subject–object quantifier interaction (like (28c) above) were produced in a manner similar to those like (28b), 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 using prosody. However, hearers' ability to retrieve the correct interpretation on the basis of prosody alone depends on the lexical items that interact and give rise to the ambiguity.

The observation of depressed percentages in the perception studies also underscores the fact that while there are clear and significant trends in hearers' ability to recruit prosody to assign meaning, what may matter most is not their OVERALL success rate, but rather the difference in responses between conditions in which one or the other interpretation is favored. Note, however, that although these findings are 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.

McMahon et al. (2004) designed four short children's stories, each of which included a sentence with ambiguous pronominal reference and an ambiguous sentence involving the interaction of a universal quantifier with negation. The story context and a continuation sentence favored one of the two scopal relations, as in (30).

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(30) Every bunny didn't jump over the fence . . .
(a) not a single one jumped over. (∀ > negation)
(b) only some did. (negation > ∀)

Eleven parents were recorded reading these stories to their children, and sentences were later coded holistically for falling/rising contour. The authors found no discernible pattern with respect to sentence-final contour, reporting a consistent 58% falling contour across all sentences.

The excised target sentences were then included in a perception study, during which hearers were told explicitly that prosody could be used to disambiguate sentences. During the test session, participants were asked to choose between two images representing what had immediately happened before a target sentence was uttered. They were given an unlimited amount of time to inspect the images 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 interpretation with negation taking wide scope 59% of the time, regardless of the interpretation intended by the speaker. This preference persisted despite the intonational contour. See Table 2.

Perceived interpretation	Rising contour	Falling contour
negation > *every*	61	58
every > negation	39	42

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

These results, then, could be taken as demonstrating that there is in fact NO connection between prosody and sentence interpretation. However, the experimental design may have 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 may also have played a significant role in that they 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. It is not possible to conclude from this study, then, that speakers do not offer the hearer such prosodic cues, and if speakers did not consistently provide prosodic cues to interpretation in this study, it is not possible to conclude that hearers cannot recruit cues that are not consistently there.

Jackson (2006) presented four speakers with nearly 200 sentences involving the scopal interaction of a small set of lexical items in different syntactic positions. Items included negation, the universal quantifier *every*, and the indefinites *a* and *a few*. Each sentence was accompanied by two still images, illustrating

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

the two interpretations (e.g. a group of circles, each of which hit a different square vs. a group of circles, all hitting the same square, for the sentence *Every circle hit a square*). Participants were asked to produce the sentence in a way that favored one of the two interpretations. 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. Productions were also not presented to hearers.

Given the previous experimental studies, the following picture emerges. First, participants of studies conducted in English and other languages CAN recruit prosodic information when disambiguating a range of sentence types. Their ability to do so, however, is dependent upon a range of factors, including the lexical items, type of ambiguity, speaker, and experimental design. Second, for the sparse evidence we have for ambiguous sentences with quantification and negation in English – precisely those sentences that have driven the debate over the years – there is still an open question regarding whether or not hearers can recruit prosodic information in the service of arriving at the correct interpretation. The current research was designed to fill this gap.

4. Stimuli from Production experiment

The auditory stimuli for the two Perception experiments reported in this paper were taken from a larger Production experiment, reported in more detail in Syrett et al. (to appear). The aim of that study was to elicit productions of ambiguous sentences in a discourse context for the purpose of finding prosodic correlates to sentential interpretation (and, therefore, scopal relation). We therefore present key aspects of the Production experiment here as a backdrop to the Perception experiments, which made use of a subset of speaker sound files.

4.1 *Experimental method*

4.1.1 *Participants*

Participants for the Production and Perception experiments were all undergraduates who received course credit in an introductory psychology or linguistics course in exchange for their participation. Nineteen native speakers of American English were included in analysis of the Production data.

4.1.2 *Stimulus design*

Test items were scopally ambiguous sentences in which negation interacted with either the universal quantifier *all* in subject position, or *many* or *most* in object position, as in (31)–(32).

(31) **All** the magnolias wo**n't** bloom. (universal quantifier, negation)
(32) Liam does**n't** know **many** alumni. (*many*, negation)

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Following the differences noted by Jackendoff (1972) and Kadmon & Roberts (1986), each test sentence type was presented in multiple discourse contexts, each favoring a different sentence interpretation and controlling for the role of negation in presupposition and assertion, and the QUD. (See Hirchberg & Avesani 1997, 2000; Baltazani 2002, 2003; and Braun 2006, among others, for similar methodology.) Sonorance of test items was controlled for in order to elicit a smooth intonational contour. There were four contexts for each of the four *all* sentences and three for each of the four *many*/*most* sentences, for a total of 28 test items.

Control items included ambiguous sentences that could be disambiguated through prosody (see Section 3 above). These items included five pairs of sentences with a *because* clause interacting with negation (as in (33)), five pairs of sentences containing a focus-sensitive operator (as in (34)–(35)) (three with *only* and two with *even*), and four pairs containing ambiguous pronominal reference (as in (36)), for a total of 14 pairs, or 28 control items.

(33) Georgia is**n't** singing **because** she's preparing for an audition. (negation, *because*)
(34) Mary **only** ran one mile. (focus sensitive operator)
(35) She **even** painted the garage. (focus sensitive operator)
(36) Alan punched Owen and then **he** kicked **him**. (pronominal reference)

As with the test items, each sentence was presented in two different contexts, each one favoring one of the two competing interpretations.

Test items were balanced between blocks, with one or two instances of a test sentence at most in each block. All items were then pseudorandomized within each block, with members of a minimal pair separated between two blocks. Items appearing in the Production experiment that were also used in the two Perception experiments are presented in Appendix A.

4.1.3 *Procedure and analysis*

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. Participants were encouraged to read the items as naturally as possible, as though they were recording them for an audiobook or reading to children.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Target sentences were excised from the surrounding context using Praat speech analysis software (Boersma 2001). They were then annotated with segments delineated from the onset and offset of each lexical item. We conducted two main analyses on the test items. First, each file was coded blindly for the type of sentence-final contour that was observed. Second, we also conducted an 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. We excluded from analysis tokens that had a comprehension score of 0, as well as 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), as were item sets that had consistently low comprehension scores or for which participants reported experiencing difficulty. Further details for all control and test items are reported in Syrett et al. (to appear).

4.2 *Main findings*

A blind coding of the sentence-final contour as falling or non-falling revealed a high percentage of falling contours across the board. Sentences in which the universal quantifier *all* interacted with negation were more likely than chance to display a falling contour, regardless of the discourse context and scopal relation. Sentences in which *many* or *most* interacted with negation displayed a falling contour more often than a non-falling contour. However, with these sentences, the percentage of falling contour varied with the scopal relation, with the highest percentage of falling contour observed when *many* or *most* took scope over negation.

We also conducted an analysis of certain acoustic information to further determine whether there were additional auditory correlates to interpretation, and found an effect of word duration. For sentences with *all* and negation, the length of the final word was longest in contexts favoring an interpretation where the quantifier took wide scope, and shortest when the context favored an interpretation in which negation took wide scope. For sentences in which *many* or *most* interacted with negation, the quantifier was in object position. The duration of this quantifier was shortest in contexts in which the quantifier took scope over negation. In this same context – as with the *all* sentences – the final word was longest. Thus, for both sets of sentences, word duration was a cue to scopal relation (although not consistently), but only for the *many*/*most* sentences was the distinction between rising and fall-rise contour indicative of scopal relation, in the direction expected. These findings are in many ways consistent with those of Albritton, McKoon & Ratcliff (1996), in which speakers did not consistently prosodically differentiate between competing versions of syntactically ambiguous sentences.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

However, it should be noted that although this was the case, and there was considerable variability among items and speakers, there were some speakers for whom the trends in production patterned closely after those discussed in the literature, and who consistently distinguished between the different interpretations in their productions. We turned to those speakers for the sound files used in our Perception experiments. Examples of their productions of minimal pair members are presented in Appendix B.

Given the previous research reviewed earlier and our Production experiment as a backdrop, we ask an important question: When 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 experiments, 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. Perception Experiment 1

The joint purpose of the two Perception experiments was to determine whether hearers can use surface level auditory cues provided by speakers to arrive at the intended interpretation of scopally ambiguous sentences. To this end, each of the experiments served a specific purpose. In Perception Experiment 1, we sought to determine whether hearers could use the prosodic delivery of a target sentence to properly situate it in a discourse context indicated by two possible continuation sentences. In Perception Experiment 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 subset of naïve speakers from the Production experiment and an experimenter. This experimenter was the second author: a female with extensive vocal musical training who was acquainted with the theoretical and experimental literature, and who was therefore well acquainted with the prosodic patterns claimed to be correlated with the various interpretations. The two experiments were run separately. Each experiment began with a brief training session, in which participants were introduced to the task structure as well as the possibility of disambiguating sentences with prosody, using items with ambiguous pronominal reference. Each experiment was completed in under 30 minutes.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

5.1 *Experimental method*

5.1.1 *Participants*

Forty-nine undergraduates participated in Perception Experiment 1. Data from four additional participants were excluded from the analysis, because the participants indicated that English was not their native language. One other participant was excluded due to consistently low response times (many below 150 ms).

5.1.2 *Stimuli*

Items for Perception Experiment 1 were contributed by four speakers from the Production experiment (one male and two females) and the experimenter. The naïve speakers were selected on the basis of 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 in their respective contexts, 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. For the test items, there were two with *all* and negation, and one with *many*/*most* and negation. For the control items, there were two with negation and *because*, and one with one of the focus-sensitive operators (*only* or *even*). These items are provided in Appendix A. The items 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 each block was randomized by the stimuli-presentation software. The order of the blocks was predetermined: the female experimenter, a naïve male, and the two naïve females.

5.1.3 *Procedure*

Items for both of the Perception experiments were presented in a quiet laboratory setting at either a Dell or an iMac computer using the SuperLab stimulus presentation software (Cedrus Corporation 2012) and a SuperLab response pad. Participants wore noise-cancelling headphones.

Each item in Perception Experiment 1 had the same 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 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 1. Additional examples for both Perception experiments are presented in Appendix B.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

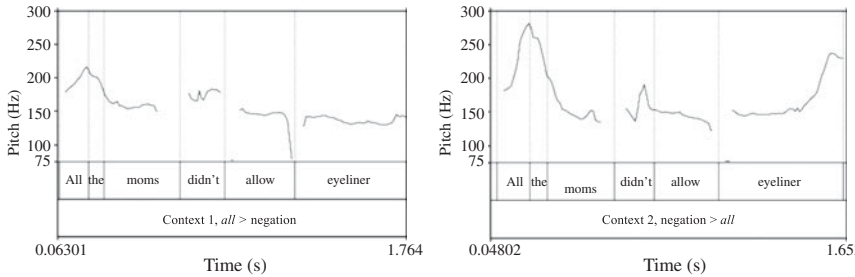


Figure 1
Pitch tracks for two versions of the sentence *All the moms didn't allow eyeliner* by Speaker 1 from two different contexts, each favoring a different scopal relation.

The auditory stimulus was presented 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 fact that the speaker had responded correctly to the comprehension question when uttering these sentences as part of the contexts in the Production experiment. Perception participants had a maximum of 10 seconds to make their decision. If no response was made during that time, the trial would end, and the next one would begin. Most speakers rendered a response well before this time. There were only 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.

5.2 *Results*

We first looked at the percentage of correct responses for each test item type, overall and by scopal relation. These results are presented in Table 3. The percentages observed here may seem somewhat low overall, given much higher success rates observed in other psycholinguistic studies, such as lexical decision or discrimination tasks.[2] To explain this rate, we first refer to the results of previous

[2] The reader may notice an egregious departure from all other averages with Speaker 3's production of a *many/most* > negation sentence (29.6%). This item appears to have been the result of hearers overwhelmingly interpreting this speaker's production of this item as having the reverse scope. Upon further inspection of this file, we discovered that this speaker appeared to place contrastive focus on the word *enjoy* in the sentence (i.e., *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 > *many/most*), this speaker uttered the sentence with a clear fall-rise contour. Accordingly, hearers responded correctly over 80% of the time.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Item type	% correct (SE)	Speaker 1 (exp'r, F)	Speaker 2 (382, M)	Speaker 3 (369, F)	Speaker 4 (362, F)
Total	67.2 (.02)	74.7	64.1	61.9	67.9
all	65.2 (.02)	72.2	62.5	68.8	56.3
all > negation	63.9 (.03)	61.4	67.1	65.9	60.2
negation > *all*	66.2 (.03)	83.0	58.0	71.6	52.3
many/most (*M*)	69.3 (.02)	77.3	64.8	55.7	79.6
M > negation	62.5 (.04)	77.3	54.6	29.6	88.6
negation > *M*	76.1 (.03)	77.3	75.0	81.8	70.5

Table 3
Average overall percentage correct in per item type and per speaker in Perception Experiment 1.

studies reviewed in the experimental background section, which report comparable rates. Indeed, these low percentages are not specific to the ambiguous quantificational sentences. Second, as a point of comparison, we note that the percentage of correct responses for the sentences with negation and a *because* clause were similar, with an overall percentage correct of 69.3%, and 62.4% when the *because* clause was intended to take wide scope, and 76.0% when negation took wide scope over the *because* clause.

To determine whether the average scores for each item type and each corresponding scopal relation were above chance, we ran a binomial probability test on the overall average (the far left column of Table 3) ($p = \cdot 5$, 99% Confidence Interval (CI)). Indeed, despite the lower-than-anticipated percentages, the averages were significantly above chance level: *all* > negation: 2.7, $p < .01$; negation > *all*: 3.1, $p = .002$; *many/most* > negation: 2.5, $p = .01$; negation > *many/most*: 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*: $\chi^2(4) = 31.45$, $p < .01$.

Admitting the possibility that there could be a response bias contributing to these scores, we ran an additional d′ analysis using signal detection theory (Macmillan & Creelman 2005), which is a measure of hearer sensitivity. Hit rates and false alarm rates were used to calculate a d′ score for each participant for the test items (d′ = (z(Hits)–z(False Alarms)). Scores can range from 0 (no detectability) to roughly 4. The average d′ score was 1.02, which was significantly above chance (0) ($t(43) = 8.82$, $p < .0001$). The hit and false alarm rates are plotted in Figure 2. Importantly, hit rates are higher than false alarm rates.

Following the experimental session, participants were asked to comment on their strategy for responding to the items, and whether any aspect of the items

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

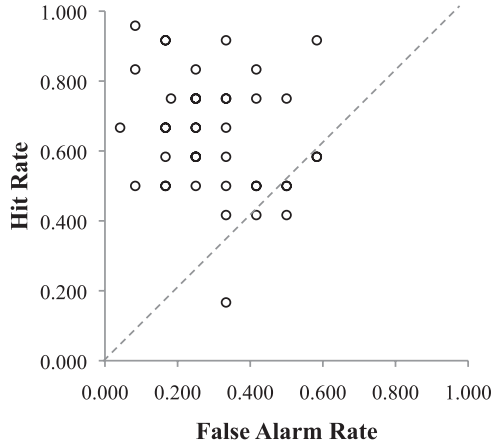


Figure 2
Hit and false alarm rates for participants' responses to test items in Perception Experiment 1.

themselves made making a decision either easier or harder (following an approach used successfully by Syrett & Kawahara, published online 14 November 2013). 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 regard to specific words), *tone*(*s*) (on words), *emphasis*, *stress*, *loud*, and *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*), and *break*(*s*). This coding method allowed for participants' responses to be coded as belonging to one or more categories. Only three of the participants' responses could not be categorized, because their responses were too vague. Two participants failed to complete a survey. We present the distribution of participant responses in Table 4.

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

Table 4
Participants' categorized responses to the post-experiment survey regarding their response strategies in Perception Experiment 1.

The responses indicate that when hearers were listening to the sentences to select a plausible continuation sentence, they were attending to those aspects

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

of the sentences that speakers were prosodically manipulating to favor one interpretation over another. It is also interesting that most hearers did NOT indicate that they attended to whether the sentence was rising or falling, suggesting that there were cues beyond sentence-final contour enabling them to make a decision.

5.3 *Discussion*

The results from Perception Experiment 1 demonstrate that when speakers prosodically favor an interpretation of a scopally ambiguous sentence supported by a discourse context, hearers can successfully recruit this auditory information to assign the correct interpretation. The differences observed among speakers and items further illustrate 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 cues encoded in the speakers' 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. 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.

6. Perception Experiment 2

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

6.1 *Experimental method*

6.1.1 *Participants*

Thirty-seven undergraduates participated in Perception Experiment 2. Ten 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 Perception Experiment 1 (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. Since this experiment asked participants to listen to a contrast between

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

two members of a minimal pair and choose the best version, given the preceding context, we decided to use those sound files that would give them the best chance of success. Each speaker provided the same six minimal pairs of sentences used in Perception Experiment 1. 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 same 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 (37). The target sentence was indicated with <<>>.

(37) 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 aurally presented with two versions of the target sentence sequentially, each accompanied by the option 'A' or 'B' on the screen. Version A was always a falling contour, and B a non-falling contour. A third screen then appeared with A and B both on it, and participants made their selection of the best match, given the preceding discourse context.

Responses registered after 3.5 seconds were not included in the analysis. We used this timeframe, because the unlike in Perception Experiment 1, in which participants were asked to choose between two sentences, in Perception Experiment 2, 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 the task and verified that 3.5 seconds was an adequate amount of time to make the choice.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

6.2 *Results*

As in Perception Experiment 1, we analyzed the percentage of correct responses for each item type, both overall and by scopal relation. The results are presented in Table 5.

Item type	% correct (SE)	Speaker 1 (exp'r, F)	Speaker 3 (369, F)	Speaker 4 (362, F)
Total	75.8 (.02)	74.5	78.5	74.3
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* (*M*)	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

Table 5
Average overall percentage correct per item type and per speaker in Perception Experiment 2.

In this experiment, 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. As before, we point out that responses to control sentences in which a *because* clause and negation scopally interacted exhibited comparable percentages: 82.4% correct overall, with 79.3% correct when the *because* clause took wide scope and 85.4% when negation took wide scope.

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, $p < .00001$. 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 also 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*: $\chi^2(3) = 43.54$, $p < .0001$.

As in Perception Experiment 1, we calculated a d′ score for each participant for the test items. The average d′ score was 1.26, which was significantly above chance (0) ($t(36) = 8.09$, $p < .0001$). The hit and false alarm rates are plotted in Figure 3. As before, hit rates are higher than false alarm rates.

The results demonstrate that participants were not choosing at random; nor were they biased towards one prosodic contour over another. Instead, their choice

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

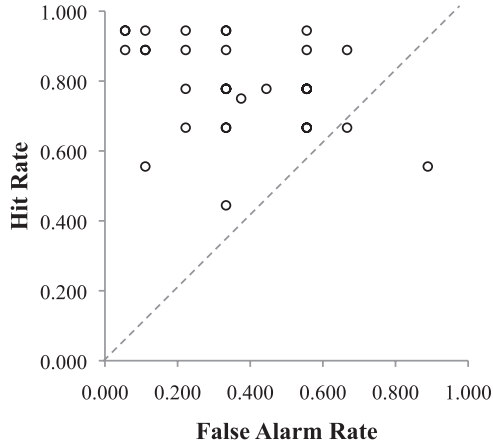


Figure 3
Hit and false alarm rates for participants' responses to test items in Perception Experiment 2.

of speaker rendition was based on the information provided in the previous context, which set up a particular information structure and a particular scopal relation.

6.3 *Discussion*

The findings of Perception Experiment 2 complement those from Perception Experiment 1 to 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 Perception experiment, participants recruited the information they received 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 successful in this task, with the notable exception of the *all* sentences, particularly where negation took wide scope, where hearers were largely at chance. It cannot be that all sentences with negation taking wide scope caused them difficulty, as they were quite successful with the *many*/*most* sentences and the negation–*because* sentences. Indeed, in these cases, hearers were highly successful with both scopal relations. What these 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),

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

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 in a position where it is 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. For now, we take the differences in response patterns as evidence that it is important to test a range of lexical items and syntactic positions when investigating similar sentences.

7. General 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 delivery. We first identified a set of target sentences involving the interaction of quantification and negation from a controlled Production experiment in which the sentences were embedded in a disambiguating discourse context and speakers prosodically differentiated between members of a minimal pair in a manner largely consistent with claims from the previous literature. We then asked whether – given these clear differences correlated with sentence-level meaning – hearers could use such information in the speech signal to arrive upon the correct, intended interpretation. In two Perception experiments, we provided an affirmative answer to this question.

In Perception Experiment 1, we asked whether hearers could use prosody to situate a target sentence 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. In Perception Experiment 2, we asked 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 rendition of the sentence that best corresponded to that interpretation. The only notable exception to participants' 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.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Thus, we conclude that when speakers provide them, hearers CAN recruit auditory cues to disambiguate scopally ambiguous sentences in precisely those cases where these 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.

As we mentioned in the introduction to this paper, 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 (38).

(38) **Every** horse did**n't** jump over the fence.

Musolino's (1998) initial observation was that children for whatever reason seem to be locked into the reading where negation takes narrow scope (i.e., the 'surface scope' reading) and have great difficulty accessing the 'inverse scope' reading where negation takes wide scope (his 'observation of isomorphism'). This pattern has been replicated across a variety of languages by a number of different researchers, using lexical items beyond the universal quantifier in subject position (see 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, and that it is indeed within their grammatical capacity (see Musolino & Lidz 2006; Gualmini et al. 2008; Viau, Lidz & Musolino 2010). Thus, they simply display an overwhelming preference for the surface scope reading, but are in principle able to access both readings. One might ask, given the current results, what role prosody has to play in children's performance. Musolino & Lidz (2006: fn. 13) remark 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 this paper going to press, this appears to still be the case.

Previous authors have indicated 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: 130) 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'. A similar statement is echoed in Lidz & Musolino (2005: 87–88). It is far from clear, however, what the 'most naturally compatible' prosody was in those studies – or,

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

given the results of our Production experiment, what this would be, even from puppeteer to puppeteer. On a related note, Gualmini et al. (2008: fn. 14) say 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: 219) do state explicitly that their test sentences, such as (39), in a context supporting negation taking wide scope, 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*'.

(39) **Every** letter was**n't** delivered.

Why they describe the 'required' intonation this way is unclear, given (i) the claims in the theoretical literature, which do not describe the required or favored prosody in this way, and (ii) the lack of experimental evidence to this effect at the time that their paper was written. This point aside, children in their experiment did access the interpretation where negation took wide scope over the quantifier 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: fn. 10). However, in Section 3, we reviewed reasons why we think that 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 auditory cues associated with quantifier–negation sentences, as the adult participants in our perception studies were.

Building upon a discussion in Musolino & Lidz (2006: 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, such as (40), may have provided children with the necessary contrast in scalar alternatives, coupled with fall-rise prosody needed to access the reading where negation took wide scope.

(40) **Every** horse jumped over the log [and/but] **every** horse did**n't** jump over the fence.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Given that 'log' and 'fence' form members of a scale (i.e., 'things that the horses could have jumped over'), both should be instantiated with fall-rise prosody – especially given the continuation rise on *log* (see 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 (see Jackendoff 1972, Kadmon & Roberts 1986) – precisely the scenario that favors fall-rise and negation taking wide scope. Of course, the findings of Viau et al. (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 earlier.

There is also promising independent evidence that children can, in fact, recruit prosodic information to calculate the partitive scalar implicature associated with *some* (Miller et al. 2005), to retrieve information about pronominal reference (Maratsos 1973), to resolve attachment ambiguities (Snedeker & Yuan 2008), and to assign syntactic structure (Beach, Katz & Skowronski 1996; Syrett 2010). 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' 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 demonstrates that part of becoming an adult language user involves acquiring the ability to recruit prosodic cues to access the interpretations of scopally ambiguous sentences in a discourse context. The findings of the complementary Perception experiments reported in this paper underscore the importance of attending to the manner in which scopally ambiguous sentences are presented in the course of psycholinguistic experiments, as hearers do demonstrate sensitivity to the relation between prosodic rendition of these sentences, the information structure of the discourse context, and the scopal relation.

APPENDIX A

Contexts and target sentences appearing in Production and Perception experiments

The following are items (contexts and target sentences) from the Production experiment used in the Perception experiments. Items used in the training session of the Perception experiments are indicated. Comprehension question in the Perception experiments follow the contexts. Discourses were minimally altered for Perception Experiment 2 to fit on the computer screen, while maintaining the integrity of the discourse context.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Test sentences: Universal quantifier *all* and negation

(1) *All the magnolias won't bloom.*

(a) Context 1: *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) Context 2: 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) Comprehension question
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.*

(a) Context 1: *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) Context 2: 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.

(c) Comprehension question
Which sentence should follow?
a. The coat was in perfect condition.
b. Only the sleeves needed repair.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(3) *All the moms didn't allow eyeliner*.

(a) Context 1: *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) Context 2: 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.

(c) Comprehension question
Which sentence should follow?
a. They were all in agreement.
b. Only the moms of the older girls let their daughters wear it.

Test sentences: *many/most* and negation

(1) *Liam doesn't know many alumni*.

(a) Context 1: *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) Context 2: 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) Comprehension question
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*.

(a) Context 1: *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.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(b) Context 2: 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) Comprehension question
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.

Control items: *because* clause and negation

(1) *Georgia isn't singing because she is preparing for an audition.*
(a) Context 1: *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) Context 2: 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) Comprehension question
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.*
(a) Context 1: *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.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(b) Context 2: 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) Comprehension question
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.*

(a) Context 1: *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) Context 2: 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) Comprehension question
Which sentence should follow?
a. It's been too cold and wet lately.
b. He just loves to do kickboxing.

Control items: focus-sensitive item *only*

(1) *Larry only elbowed Riley.*

(a) Context 1: 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.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(b) Context 2: 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) Comprehension question
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.*

(a) Context 1: 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) Context 2: focus on DP (numeral)
This morning, Mary (who used to run every day, but had not done so in years) decided to attempt a five-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 five miles that morning, but her accomplishment was much more modest. Mary only ran one mile. She should have run for longer.

(c) Comprehension question
What sentence would most logically come next?
a. She should have also swam and biked.
b. She should have run for longer.

Control items: focus-sensitive item *even*

(1) *She even painted the garage.*

(a) Context 1: 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.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(b) Context 2: 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) Comprehension question
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*.

(a) Context 1: 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) Context 2: 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) Comprehension question
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: pronominal reference

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

(a) Context 1: 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) Context 2: 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.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

(c) Comprehension question
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 item)
(a) Context 1: 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) Context 2: 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) Comprehension question
Who drove off the road?
a. Ryan.
b. Nolan.

(3) *Mary admires Arianna but she doesn't like her.*
(training item)
(a) Context 1: 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) Context 2: 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) Comprehension question
Which is true?
a. Mary doesn't like Arianna.
b. Arianna doesn't like Mary.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

APPENDIX B

Pitch tracks representing minimal pair members for target sentences presented to hearers in Perception Experiments 1 and 2

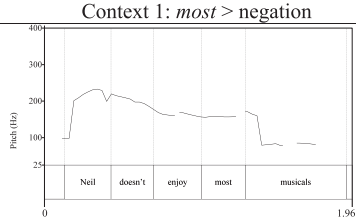
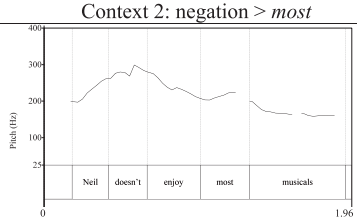
Context 1: *most* > negation	Context 2: negation > *most*
	
% correct: Perception 1: 88.6 Perception 2: 91.7	% correct: Perception 1: 70.5 Perception 2: 91.9

Figure B1
Target sentence: *Neil doesn't enjoy most musicals*; Speaker 4.

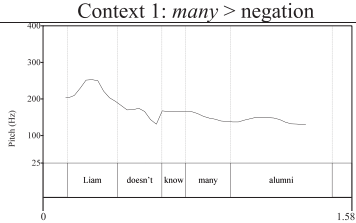
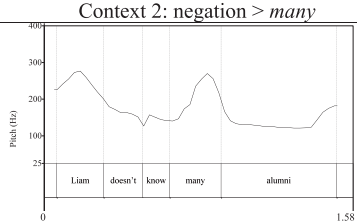
Context 1: *many* > negation	Context 2: negation > *many*
	
% correct: Perception 1: 77.3 Perception 2: 88.9	% correct: Perception 1: 77.3 Perception 2: 66.7

Figure B2
Target sentence: *Liam doesn't know many alumni*; Speaker 1.

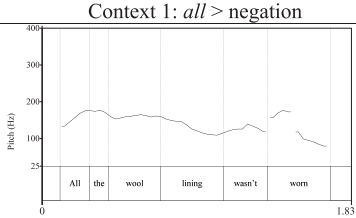
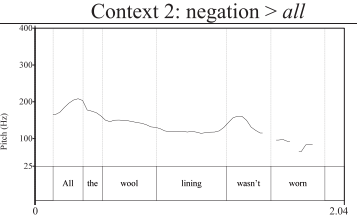
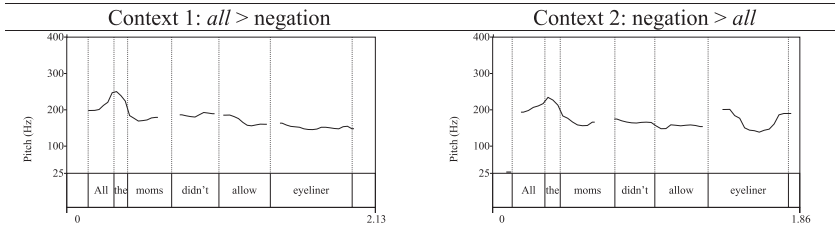
Context 1: *all* > negation	Context 2: negation > *all*
	
% correct: Perception 1: 70.5 Perception 2: n/a	% correct: Perception 1: 81.8 Perception 2: n/a

Figure B3
Target sentence: *All the wool lining wasn't worn*; Speaker 2.

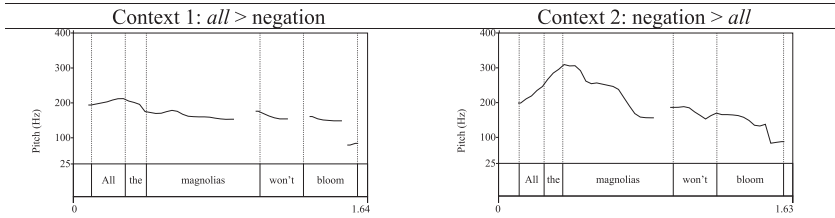
https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press



% correct:
Perception 1: 68.2 Perception 2: 82.4

% correct:
Perception 1: 77.3 Perception 2: 38.2

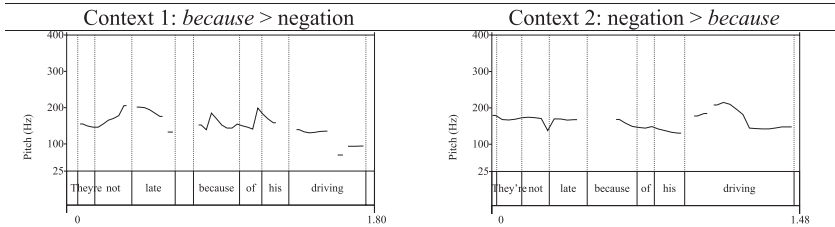
Figure B4
Target sentence: *All the moms didn't allow eyeliner*; Speaker 3.



% correct:
Perception 1: 65.9 Perception 2: 64.7

% correct:
Perception 1: 65.9 Perception 2: 76.5

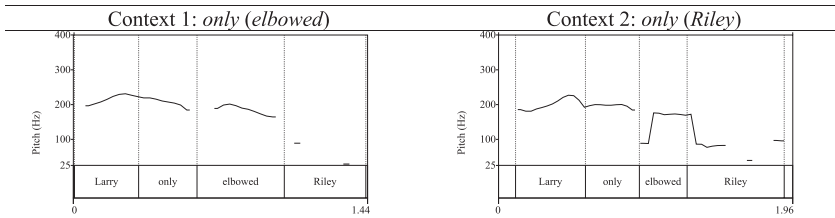
Figure B5
Target sentence: *All the magnolias won't bloom*; Speaker 3.



% correct:
Perception 1: 59.1 Perception 2: 88.6

% correct:
Perception 1: 81.4 Perception 2: 73.3

Figure B6
Target sentence: *They're not late because of his driving*; Speaker 1.



% correct:
Perception 1: 79.5 Perception 2: 80.6

% correct:
Perception 1: 82.4 Perception 2: 85.7

Figure B7
Target sentence: *Larry only elbowed Riley*; Speaker 4.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

REFERENCES

Akmajian, Adrian & Ray S. Jackendoff. 1970. Coreferentiality and stress. *Linguistic Inquiry* 1.1, 124–126.

Albritton, David W., Gail McKoon & Roger Ratcliff. 1996. Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22.3, 714–735.

Baltazani, Mary. 2002. The prosodic structure of quantificational sentences in Greek. In Mary Andronis, Erin Debenport, Anne Pycha & Keiko Yoshimura (eds.), *Chicago Linguistic Society* (CLS) 38, 63–78. Chicago, IL: Chicago Linguistic Society.

Baltazani, Mary. 2003. *Quantifier scope and the role of intonation in Greek*. Ph.D. dissertation, University of California, Los Angeles.

Beach, Cheryl M. 1991. The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language* 30.6, 644–663.

Beach, Cheryl M., William F. Katz & Aire Skowronski. 1996. Children's processing of prosodic cues for phrasal interpretation. *Journal of the Acoustical Society of America* 99.2, 1148–1160.

Bel, Bernard & Isabelle Marlien (eds.). 2002. *Speech Prosody 2002*. Aix-en-Provence: Laboratoire Parole et Langage.

Boersma, Paul. 2001. Praat: A system for doing phonetics by computer. *Glot International* 5.9–10, 341–345.

Bolinger, Dwight L. 1965. *Forms of English: Accent, morpheme, order*. Edited by Isamu Abe & Tetsuya Kanekiyo. Cambridge, MA: Harvard University Press.

Braun, Bettina. 2006. Phonetics and phonology of thematic contrast in German. *Language and Speech* 49.4, 451–493.

Braun, Bettina & Aoju Chen. 2010. Intonation of 'now' in resolving scope ambiguity in English and Dutch. *Journal of Phonetics* 38.3, 431–444.

Büring, Daniel. 1997. The great scope inversion conspiracy. *Linguistics and Philosophy* 20.2, 175–194.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics and Philosophy* 26.5, 511–545.

Cedrus Corporation. 2012. SuperLab 4.5 [Stimulus presentation software]. San Pedro, CA.

Cooper, William E. & Jeanne Paccia-Cooper. 1980. *Syntax and speech*. Cambridge, MA: Harvard University Press.

Fodor, Janet Dean. 2002a. Prosodic disambiguation in silent reading. In Masako Hirotani (ed.), *North East Linguistic Society* (NELS) 32, 113–132. Amherst, MA: GLSA.

Fodor, Janet Dean. 2002b. Psycholinguistics cannot escape prosody. In Bel & Marlien (eds.), 83–90.

Frazier, Lyn & Charles Clifton Jr. 1996. *Construal*. Cambridge, MA: MIT Press.

Gualmini, Andrea, Sarah Hulsey, Valentine Hacquard & Danny Fox. 2008. The question–answer requirement for scope assignment. *Natural Language Semantics* 16.3, 205–237.

Hirschberg, Julia. 1985. *A theory of scalar implicature (natural languages, pragmatics, inference)*. Ph.D. dissertation, University of Pennsylvania.

Hirschberg, Julia & Cinzia Avesani. 1997. The role of prosody in disambiguating potentially ambiguous utterances in English and Italian. In Antonis Botinis, Georgios Kouroupetroglou & George Carayiannis (eds.), *Intonation: Theory, models, and applications* [Proceedings of European Speech Communication Association (ESCA) Workshop], 189–192. European Speech Communication Association & University of Athens Department of Informatics.

Hirschberg, Julia & Cinzia Avesani. 2000. Prosodic disambiguation in English and Italian. In Antonis Botinis (ed.), *Intonation: Theory, models, and applications*, 87–95. Dordrecht: Kluwer.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jackson, Scott. 2006. Prosodic phrasing and logical scope in English. Presented at the 19th Annual CUNY Conference on Human Sentence Processing, New York.

Jacobs, Joachim. 1984. Funktionale Satzperspektive und Illokutionssemantik. *Linguistische Berichte* 91, 25–58.

Kadmon, Nirit & Craige Roberts. 1986. Prosody and scope: The role of discourse structure. In Anne M. Farley, Peter T. Farley & Karl-Erik McCullough (eds.), *Parasession on Pragmatics and Grammatical Theory: Chicago Linguistics Society* (CLS) 22, Part 2, 16–28. Chicago, IL: Chicago Linguistic Society.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Kitagawa, Yoshihisa & Janet Dean Fodor. 2006. Prosodic influence on syntactic judgments. In Gisbert Fenselow, Caroline Fery, Matthias Schlesewsky & Ralf Vogel (eds.), *Gradience in grammar: Generative perspectives*, 336–358. Oxford: Oxford University Press.
Koizumi, Yukiko. 2009. *Processing the not-because ambiguity in English: The role of pragmatics and prosody*. Ph.D. dissertation, City University of New York.
Krifka, Manfred. 1998. Scope inversion under the rise-fall contour in German. *Linguistic Inquiry* 29.1, 75–112.
Ladd, D. Robert Jr. 1980. *Structure of intonational meaning: Evidence from English*. Bloomington, IN: Indiana University Press.
Ladd, D. Robert Jr. 1986/2008. *Intonational phonology*. Cambridge: Cambridge University Press.
Lehiste, Ilse. 1973. Phonetic disambiguation of syntactic ambiguity. *Glossa: An International Journal of Linguistics* 7.2, 107–122.
Lehiste, Ilse, Joseph P. Olive & Lynn A. Streeter. 1976. Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America* 60.5, 1199–1202.
Liberman, Mark & Ivan A. Sag. 1974. Prosodic form and discourse function. In Michael W. LaGaly, Robert A. Fox & Anthony Bruck (eds.), *Chicago Linguistics Society* (CLS) 10, 416–427. Chicago, IL: Chicago Linguistic Society.
Lidz, Jeffrey & Julien Musolino. 2002. Children's command of quantification. *Cognition* 84.2, 113–154.
Lidz, Jeffrey & Julien Musolino. 2005. On the status of quantificational indefinites: The view from child language. *Language Acquisition* 13.2, 73–102.
Macmillan, Neil A. & C. Douglas Creelman. 2005. *Detection theory: A user's guide*. Mahwah, NJ: Lawrence Erlbaum.
Maratsos, Michael. 1973. The effects of stress on the understanding of pronominal co-reference in children. *Journal of Psycholinguistic Research* 2.1, 1–8.
Martí, Luisa. 2001. Intonation, scope, and restrictions on quantifiers. In Karine Megerdoomian & Leora Anne Bar-el (eds.), *West Coast Conference on Formal Linguistics* (WCCFL) 20, 372–385. Somerville, MA: Cascadilla Press.
McMahon, Erin, Jeffrey Lidz & Janet Pierrehumbert. 2004. Suprasegmental cues to meaning in child-directed speech. Presented at the 17th CUNY Sentence Processing Conference, College Park, MD.
Miller, Karen, Cristina Schmitt, Hsiang-Hua Chang & Alan Munn. 2005. Young children understand *some* implicatures. In Alejna Brugos, Manuella R. Clark-Cotton & Seungwan Ha (eds.), *Boston University Conference on Language Development* (BUCLD) 29.2, 389–400. Somerville, MA: Cascadilla Press.
Musolino, Julien. 1998. *Universal Grammar and the acquisition of semantic knowledge: An experimental investigation into the acquisition of quantifier–negation interaction in English*. Ph.D. dissertation, University of Maryland, College Park.
Musolino, Julien. 2011. Studying language acquisition through the prism of isomorphism. In Jill de Villiers & Tom Roeper (eds.), *Handbook of generative approaches to language acquisition*, 319–349. New York: Springer Science & Business Media.
Musolino, Julien & Jeffrey Lidz. 2006. Why children aren't universally successful with quantification. *Linguistics* 44.4, 817–852.
Nagel, H. Nicholas, Lewis P. Shapiro & Rebecca Nawy. 1994. Prosody and the processing of filler–gap sentences. *Journal of Psycholinguistic Research* 23.6, 473–485.
Nagel, H. Nicholas, Lewis P. Shapiro, Betty Tuller & Rebecca Nawy. 1996. Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing. *Journal of Psycholinguistic Research* 25.2, 319–344.
Price, Patti J., Mari Ostendorf, Stefanie Shattuck-Hufnagel & Cynthia Fong. 1991. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90.6, 2956–2970.
Roberts, Craige. 1996. Information structure: Towards an integrated formal theory of pragmatics. In Jae Hak Yoon & Andreas Kathol (eds.), *Papers in Semantics* (Ohio State University Working Papers in Linguistics 49), 91–136. Columbus, OH: The Ohio State University Department of Linguistics.
Roland, Douglas, Hongoak Yun, Jean-Pierre Koenig & Gail Mauner. 2012. Semantic similarity, predictability, and models of sentence processing. *Cognition* 122.3, 267–279.

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press

Sauerland, Uli & Oliver Bott. 2002. Prosody and scope in German inverse linking constructions. In Bel & Marlien (eds.), 623–628.
Schäfer, Martin. 2004. Manner adverbs and scope. In Anita Steube (ed.), *Grammatik und Kontext: Zur Interaktion von Syntax, Semantik und Prosodie bei der Informationsstrukturierung* (Linguistische Arbeitsberichte 81), 33–51. Leipzig: Institut für Linuistik Universität.
Snedeker, Jesse & Sylvia Yuan. 2008. Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language* 58.2, 574–608.
Speer, Shari R., Robert G. Crowder & Lisa M. Thomas. 1993. Prosodic structure and sentence recognition. *Journal of Memory and Language* 32.3, 336–358.
Syrett, Kristen. 2010. The representation and processing of measure phrases by four-year-olds. In Katie Franich, Kate Iserman & Lauren L. Keil (eds.), *Boston University Conference on Language Development* (BUCLD) 34, 421–432. Somerville, MA: Cascadilla Press.
Syrett, Kristen & Shigeto Kawahara. Production and perception of listener-oriented clear speech in child language. *Journal of Child Language*, doi:10.1017/S0305000913000482. Published online by Cambridge University Press, 14 November 2013.
Syrett, Kristen, Georgia Simon & Kirsten Nisula. To appear. Prosodic disambiguation of scopally ambiguous sentences. *North East Linguistic Society* (NELS) 43.
Viau, Joshua, Jeffrey Lidz & Julien Musolino. 2010. Priming of abstract logical representations in 4-year-olds. *Language Acquisition* 17.1, 26–50.
Ward, Gregory & Julia Hirschberg. 1985. Implicating uncertainty: The pragmatics of Fall-Rise intonation. *Language* 61.4, 747–776.

Authors' address: *(Syrett)*
Department of Linguistics, Rutgers,
The State University of New Jersey – New Brunswick,
18 Seminary Place, New Brunswick, NJ 08901, USA
k-syrett@ruccs.rutgers.edu

https://doi.org/10.1017/S0022226714000012 Published online by Cambridge University Press