THE ROLE OF PROSODY IN DISAMBIGUATING POTENTIALLY AMBIGUOUS UTTERANCES IN ENGLISH AND ITALIAN

Julia Hirschberg and Cinzia Avesani

AT&T Labs–Research
180 Park Avenue, Florham Park NJ 07932-0971, USA
julia@research.att.com

Centro Studi per le Ricerche di Fonetica del CNR
Padova, Italy
avesani@SABSNS.sns.it

ABSTRACT

We investigate the role that intonation plays in disambiguating potentially ambiguous utterances in English and Italian, to see a) whether speakers employ intonational means to disambiguate these utterances, and b) whether speakers of the two languages employed consistently different intonational strategies in this disambiguation. We find that, while some semantic phenomena are consistently disambiguated by both sets of speakers, the syntactic phenomena tested are not. We suggest a possible explanation for this disparity.

1. INTRODUCTION

It is widely believed that syntactic and semantic ambiguities, such as the scope of negation and quantifiers, the association of focus-sensitive operators, and the attachment of prepositional phrases, adverbials, and relative clauses can be disambiguated intonationally [5, 4, 7]. While there has been some experimental investigation of the role prosody plays in influencing hearers' interpretation of certain syntactic ambiguities in English [3, 6], there has been little empirical study of these phenomena in other languages, and even less cross-linguistic research. In this paper we present results of a production study designed to compare the mechanisms speakers employ to disambiguate syntactically and semantically ambiguous utterances in English and Italian. We wanted to discover, first, whether phenomena believed to be intonationally disambiguable would be so disambiguated by naive subjects. Second, we wanted to see whether speakers of two languages in which prosodic features such as phrasing and pitch accent can be freely varied to convey differences in meaning would use those features similarly or not.

In an earlier production study comparing English, Italian, and Spanish [2], we found that native speakers of these languages did differentiate among some types of syntactic and scopal ambiguity intonationally, employing differences in intonational phrasing and prominence. Their strategies differed among languages, with Spanish and Italian patterning together more often than either patterned with English. English, Spanish, and Italian speakers were most similar in their disambiguation of the scope of negation, employing variation in prosodic phrasing to distinguish wide from narrow scope productions, with wide scope utterances produced as a single phrase and narrow produced as two phrases. Italian and Spanish speakers also differentiated wide from narrow scope by similar variation in phrasing; however, they also placed nuclear stress on the verb to indicate wide scope negation, while English speakers located nuclear stress later in the utterance. Also, English speakers further distinguished wide from narrow scope by utterance-final tonal variation, with continuation rise employed for wide scope readings and falling intonation for narrow. While our Italian speakers consistently used phrasing variation to indicate differences in PP attachment (between NP and VP attachment), English and Spanish subjects were inconsistent in this regard. For quantifier disambiguation, the picture was more complex: For Italian and Spanish speakers, renditions of sentences containing scope-ambiguous negative quantifiers were disambiguated by variation in nuclear stress placement and in prosodic phrasing; for two English speakers, accent placement served to disambiguate these utterances. In cases of association with focus, *only/solo/sólo* was treated inconsistently by speakers of all three languages.

2. CURRENT STUDY

In order to test the validity of our previous results, we conducted a larger study of English and Italian. We constructed materials for a production study to test intonational variation in sentences that contained the following potential ambiguities: *pp* and adverbial attachment, relative clauses, focus sensitive operators (*only* and *even*), the scope of negation, and quantifier scope (*none*). We embedded these sentences in paragraphs designed to disambiguate the potential ambigu. We embedded these sentences in paragraphs designed to disambiguate the potential ambiguity. For example, '*William isn't drinking because he's unhappy*' was embedded in paragraphs in which context favored (1a) the reading in which William drinks, but not because he is unhappy and (1b) the reading in which William does not drink, because of his unhappiness.

(1) a. William is a hopeless case. There's nothing anyone can do to make him stop drinking. For a while, his friends thought he only drank to forget his troubles. But that doesn't seem to be true. William isn't drinking because he's unhappy. He drinks because he's an alcoholic.

b. When my friend William is happy, he loves to go to parties, and he can drink more than anyone I know. But lately he never wants to come to parties. When he does, he only drinks water. I think I understand. William isn't drinking because he's unhappy. And I don't know how long this will last.

There were a total of 21 pairs of paragraphs, three for each of the phenomena under examination.

Twelve subjects (six native speakers of standard American English and six of Tuscan Italian) were given each pair of paragraphs, asked to read them over for understanding and then were recorded reading them aloud. After reading each pair, speakers were then asked to answer a series of questions designed to elicit their interpretation of the target sentences in the differing contexts. We wanted to make sure that the interpretations intended for each sentence in each context were in fact the interpretations the subjects understood and had been trying to convey in reading the paragraphs aloud. First, subjects were asked to explain in their own words the difference in meaning of the target sentence in each context. A typical answer from one subject for paragraph (1a) was "It is not the case that William's unhappiness is causing him to drink"; for (1b) the same subject wrote "William is unhappy so he isn't drinking." After completing this task, subjects were then asked to answer a forced choice question for each paragraph; the following was asked for paragraphs (1a) and (1b): "Does William drink?" Both the free and forced choice tasks were scored by one of the authors to determine whether the subject had interpreted the task as intended for the purpose of the experiment or not.

Subjects' speech productions were excised from the disambiguating contexts, pitch-tracked using Entropic WAVES software, and labeled by someone who had not heard the utterance in context, using the ToBI labeling scheme. Pitch accents, phrase accents, boundary tones, and relative prominence of accents were marked.

3. RESULTS

Preliminary analysis of our data indicate, first, that both English and Italian subjects had no difficulty in understanding the differences in interpretation that we hoped to convey with our disambiguating contexts. Nearly all explained the differences as we had intended to convey them in the free-form condition, and similarly answered the questions posed in ways that confirmed their interpretation of the paragraphs was the same as the intended interpretation. Of the six English subjects' written responses to the two conditions for each of the 42 paragraphs (N=252), only six responses indicated that the subject had understood a different interpretation of the target utterance from the intended one; four of these were from one of our speakers. For the six Italian speakers, only two responses showed a different interpretation from the intended one.

The speech productions from the English speakers show clear trends for all speakers in their production of sentences ambiguous with respect to scope of negation and the focus of the focus-sensitive operators, *only* and *even*. Results are less clear for sentences where the ambiguity derived from attachment distinctions and there are no discernible patterns in the quantifier scope productions.

The English speakers disambiguate scope of negation, in paragraphs like (a) and (b) above, primarily by varying the phrasing of the ambiguous sentence. Target utterances produced in the wide scope condition (1a) rarely (2/18) contain internal phrase boundaries, while utterances produced in the narrow scope context (1b), usually (12/18) exhibit major or minor prosodic phrase boundaries before the subordinate conjunction ('because' in all cases in our study). Wide scope productions often end in 'continuation rise' (L-H%) (10/18), while narrow scope productions usually were falling contours (L-L%) (15/18). Examples of typical productions for (1a) and (1b) are (2a) and (2b), respectively.

(2) a. H* William isn't !H* drinking because he's H* unhappy L-H%

b. H* William isn't !H* drinking L-H% because he's H* unhappy L-L%

These findings are similar to those of our earlier study [2].

These speakers also display consistency in their production of variation in the focus of *even* and *only*. Examples of such sentences include '*Harold even telegraphed the paper*', where contexts vary the focus of *even* from the verb to the direct object, and '*He only wounded Anne*', where *wounded* and *Anne* alternate as focii. In 64/72 target utterances in this category, the focus of the operator represent the nuclear stress of the utterance; 30/36 utterances with *even* exhibit this pattern and 34/36 with *only*, as illustrated in (3):

(3) a. He H* only H* wounded Anne L-L%

b. He H* only !H* wounded !H* Anne L-L%

Attachment ambiguities produced different results, depending upon syntactic category: English speakers make no clear prosodic distinctions in their production of sentences ambiguous with respect to PP attachment. So, sentences such as '*He managed to find the woman with the binoculars*' are produced similarly, whether the VP or the NP attachment of the prepositional phrase is favored by context. However, speakers do exhibit some regularities in their disambiguation of adverbial and relative clause attachment, in producing sentences such as '*He had spoken to her quite clearly*' and '*The professor who loves jelly beans died in terrible pain*'. In contexts favoring the S attachment of the adverbial expression (e.g. "It was quite clear that he had spoken to her."), speakers produce the target sentence with an internal prosodic boundary separating the adverb from the remainder of the sentence (11/18) (e.g. '*He had L+H* spoken to her L-H* quite !H* clearly L-L%*) while in contexts favoring a VP attachment (e.g. "He had spoken to her in a clear manner.") only 1/18 productions exhibit such an internal boundary. For contexts favoring a non-restrictive reading of target sentences containing a relative clause, speakers produce utterances in which the clause represented a separate prosodic phrase in 9/18 cases (e.g. '*The H* professor L-H% who loves H* jelly beans L-H% H* died in !H** *terrible H* pain L-L%*'; only four productions of the restrictive reading exhibit such phrasing. So, there is weak evidence that speakers employ phrasing to signal attachment differences for some phenomena..

Speakers' productions of sentences containing the quantifier *none* do not show any consistent distinctions for the wide vs. narrow scope contexts, despite the fact that speakers **do** distinguish between the two readings. In most cases, speakers vary the accenting of the quantifier itself; 13/18 production pairs in this category showed such differences, which ranged from deaccenting the quantifier in the narrow scope condition and accenting in the wide in a few (3) cases, to varying the type of pitch accent between the two conditions. However, there are no clear differences in choice of accent for either of the conditions (e.g. 7/18 of the narrow scope cases and 11/18 of the wide bore **L+H*** accents. Nor is there a distinction in relative prominence between the two; in 10/18 cases the quantifier receives nuclear stress in the narrow condition, and similarly in 12/18 of the wide. It might be argued that the distinction between readings might have been more difficult for speakers to comprehend than for some of the other phenomena, thus leading to less consistency in productions. Paragraphs such as (4) certainly seem more artificial than (1), for example.

(4) a. It is really hard to say who you should vote for for the town council. All of the Democratic candidates are pretty mediocre, but none of them is actually corrupt. *The election of none of these candidates would be a disaster.* Any of them would probably do a decent enough job.

b. It's absolutely essential that the Democratic party win some seats in the coming election, or nothing is ever going to get better in our state. They are the only party who can save us from bankruptcy. *The election of none of these candidates would be a disaster.* Unless at least some of them are elected we will really be in trouble.

However, only one of the speakers exhibited any interpretation problems with these sentences in the written condition. And the lack of consistency we find here is consistent with our earlier results for English speakers, reported in [2].

A systematic use of intonational phrasing appears to be the main strategy adopted by our Italian speakers in disambiguating the scope of negation. All but one instance of the narrow scope utterances (16/17)[1] is uttered in two intermediate phrases, with a high (10/16) or low (6/16) phrase accent marking the syntactic boundary between the two clauses. All instances of the wide scope reading are uttered as a single intonational phrase, with a nuclear pitch accent on the VP of the main clause and deaccenting the remainder of the utterance. The intonational contours of both the wide (5a) and the narrow scope (5b) readings end in a fall, as it is illustrated in the following example:

(5) a. H* Guglielmo non H+L* beve perché é infelice L-L%

b. H* Guglielmo non H* beve H- perché é H+L* infelice L-L%

[1] One production is missing.

The difference in the scope of focus-sensitive operators such as *solo* (*only*) and *anche* (*even*) is generally conveyed by accent placement. The alternative readings of 7 out of 36 ambiguous sentences in this category did not show any prosodic evidence of disambiguation by the speakers. Of the 58 target utterances (of 29 ambiguous pairs) that did show evidence of disambiguation, 28 display a pitch accent associated both with the focus-sensitive operator and with the focussed item. The item not in focus is deaccented. So, (6a) only Anna was wounded and no one else, and in (6b), Anna was only wounded, but not, say, killed.

(6) a. Ha H* solo ferito H+L* Anna L-L%

b. Ha H* solo H+L* ferito Anna L-L%

The focus of the operator in these cases represents the nuclear stress of the utterance. In 15/30 remaining utterances, the focussed word occurs in sentence final position and is preceded by a prenuclear accent, as in '*Ha H* anche H* abbracciato il H+L* poliziotto L-L%*'. The possible ambiguity caused by accenting both candidates for operator focus is solved by speakers attributing to the focussed item a noticeably higher degree of prominence, attained through a higher F0 and/or a longer vowel duration.

Ambiguous attachment of prepositional phrases, adverbials and relative clauses is resolved by Italian speakers with less consistency. Only 11/18 sentences with ambiguous attachment of the prepositional phrase, 13/18 sentences with ambiguous adverbial attachment and 8/18 sentences with embedded relative clauses are disambiguated prosodically. Globally, the preferred strategy used by our speakers in attachment disambiguation is to vary intonational phrasing. Among the two possible attachments a PP or ADVP and a relative clause can have in our sentences, if a phrase attaches to the syntactically higher node, it is prosodically separated from the remainder of the utterance via a H- or L- phrase boundary. (10/11 for PP attachment, 12/13 for adverbial attachment, 6/8 for relative clauses). For example, in a sentence such as *Ha provato a mettersi in contatto col suo oculista a Roma*, the PP is set off in an intermediate phrase if the intended meaning is that "the subject was in Rome when he tried to get in touch with his ophthalmologist" (VP attachment); and it is part of a larger intermediate phrase — together with the phrase it attaches to — in the case of NP (or PP) reading ("the ophthalmologist and not the subject was in Rome"). The prosodic boundary is usually marked by a constellation of different cues in addition to F0 movement. When the phrase accent is not used, the intended attachment is conveyed by relative prominence of the pitch accents across the boundary, pre-boundary final lenghthening and other sandhi phenomena (3/11). In '*Lui le aveva parlato chiaramente*', the adverbial phrase is set off as a separate intermediate phrase if the adverbial is intended to attach to S ("it was obvious that he spoke to her"), and is part of a larger phrase otherwise (VP attachment, i.e. "he spoke to her in a clear manner"). In '*la ragazza che mi ha fregato il taxi era una testimone dell'omicidio*', a L- or H- phrase boundary sets off the relative clause from the NP *the girl* when the clause has a non restrictive meaning. Despite the fact that the same prosodic phenomenon (phrasing) is used in almost all cases of attachment that have been disambiguated, only 60% of our sentences showed evidence of prosodic disambiguation.

Ambiguity in the scope of the negative quantifier (e.g., for '*La presenza di nessuno dei professori la metterebbe in imbarazzo* the readings "the absence of all the professors would embarrass her", and "there are no professors whose presence could embarrass her" is distinguished by Italian speakers in 16/18 of such sentences. Two different strategies are used. Both the quantifier and the NP bound by it (*nessuno dei professori*) are accented in one reading ((7a)) and either the quantifier is accented and the NP deaccented ((7b)) or the quantifier deaccented and the NP accented in the alternative reading.

(7) a. La H* presenza di H* nessuno dei H* professori la metterebbe in H+L* imbarazzo L-L%

b. La H* presenza di H* nessuno dei professori la metterebbe in imbarazzo L-L%

Other speakers (11/18 productions) accent both the quantifier and the bound NP in both readings, but the pitch accents differ in prominence. The two readings are distinguished by: (a) assigning more prominence (higher F0) to either the quantifier or to the noun phrase in one reading and reversing that order in the alternative one. (b) Changing the difference in prominence between the quantifier and the bound noun phrase. The difference is higher in one reading and lower in the alternative one. Only two of our speakers use one of the above strategies consistently, however.

4. CONCLUSION

Comparing the production of our English and Italian speakers, we see that, for both, semantic ambiguities were disambiguated more consistently than were syntactic ambiguities. So, the scope of negation and of focus-sensitive operators were reliably distinguished by both groups, both using phrasing to disambiguate scope of negation and accent placement to distinguish focussed items. However, another semantic ambiguity, quantifier scope, was **not** consistently disambiguated by speakers of either language. These results are consistent with our previous findings for scope of negation, for both English and Italian speakers. However, our earlier studies showed no clear patterns for the disambiguation of focus ambiguities but did find consistent behavior in Italian speakers' disambiguation of quantifier scope.

Syntactic ambiguities were in general much less clearly disambiguated by speakers of either language. PP attachment, in fact, was disambiguated consistently by neither English nor Italian speakers, while only English speakers showed clear trends in the disambiguation of ambiguously attached adverbials and results for ambiguous relative clause attachment is quite mixed. However, in all cases, speakers who **did** consistently disambiguate syntactic ambiguities did so by varying intonational phrasing.

While much has been proposed for the role of prosodic phenomena as contributing additional information essential for the processing of syntactic and semantic constructions, our findings suggest that this role is not a simple one. Phenomena such as attachment decisions, in particular, which seem intuitively easy to disambiguate intonationally, rarely **were** so disambiguated, in context, even by subjects who understand their ambiguity, in contrast to phenomena such as focus location and scope of negation. One possible explanation for this is that our 'disambiguating contexts' were, in some cases, too successful. Often phenomena disambiguable by prosodic variation have a 'neutral' (or 'unmarked') production, which is felicitous for either interpretation. For example, sentences in which attachment may be signalled by the variation of internal phrase boundaries may usually be produced without internal boundaries in either condition, if the ambiguity is resolvable by other means. Thus the disambiguating contexts in which target sentences were embedded may in fact have allowed subjects felicitously to produce intonationally ambiguous utterances. While this explanation might account for the lack of prosodic distinctions for some sentences, however, it would not explain those productions of sentences also in clearly disambiguating contexts which **did** exhibit consistent intonational variation, except by the hypothesis that, in such cases, a 'neutral' production was less available.

References

[1] B. Altenberg. *Prosodic Patterns in Spoken English: Studies in the Correlation between Prosody and Grammar for Text-to-Speech Conversion*, volume 76 of *Lund Studies in English*. Lund University Press, Lund, 1987.

[2] C. Avesani, J. Hirschberg, and P. Prieto. The intonational disambiguation of potentially ambiguous utterances in English, Italian, and Spanish. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, volume 1, pages 174–177, Stockholm, August 1995.

[3] C. Beach. The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language*, 30:644–663, 1991.

[4] D. Bolinger. *Intonation and Its Uses: Melody in Grammar and Discourse*. Edward Arnold, London, 1989.

[5] R. S. Jackendoff. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge MA, 1972.

[6] P. Price, S. Shattuck-Hufnagel, and C. Fong. The use of prosody in syntactic disambiguation. *Journal of the Acoustic Society of America*, 90(6):2956–2970, 1991.

[7] L. Renzi, editor. *Grande grammatica di consultazione*, volume 1. Il Mulino, Bologna, 1988.