The Predominance of Nonstructural Factors in the Processing of Gapping Sentences

John C. J. Hoeks (j.c.j.hoeks@rug.nl)
Center for Language and Cognition, University of Groningen,
PO Box 716, 9700 AS Groningen, The Netherlands.

Petra Hendriks (p.hendriks@rug.nl)
Center for Language and Cognition, University of Groningen,
PO Box 716, 9700 AS Groningen, The Netherlands.

Louisa J. Zijlstra (l.j.zijlstra.1@student.rug.nl)
Center for Language and Cognition, University of Groningen,
PO Box 716, 9700 AS Groningen, The Netherlands.

Abstract

Two studies investigated the effects of pragmatic context and prosody on the processing of sentences like *John greeted Paul yesterday and Ben today*. This sentence is ambiguous between the clearly preferred 'nongapping' reading, where *John greeted Ben*, and the unpreferred 'gapping' reading, where *Ben greeted Paul*. Participants listen to dialogues and give a speeded response as to which reading of an ambiguous target sentence first comes to mind. The results show that context and prosody have independent and strong effects on the choice for gapping, and that, in the right combination, they can make gapping the preferred reading. Thus, even if structural factors were to play a role in processing gapping, they are outweighed by context and prosody.

Introduction

The phenomenon of 'gapping' has long been a thorn in the side of formal and psychological theories of language, mainly because it involves the seemingly unprincipled omission of words from an utterance without changing its formal meaning. Linguistic theory has difficulty assigning syntactic structure to words or phrases that are not actually present. For psycholinguistic models, modelling the comprehension of gapping is also problematic because it crucially hinges on the use of prosodic information, a factor which, due to its multi-dimensionality and complexity, is still poorly understood. Consider sentence (1a).

1a. John greeted Paul yesterday and Ben today.
1b. John greeted Paul yesterday and Ben ~~greeted Paul~~ today.
1c. John greeted Paul yesterday and ~~John greeted~~ Ben today.

In this sentence it is impossible to *uniquely* identify which elements were left out; the sentence is ambiguous between reading (1b), where first John greets Paul, and then Ben greets Paul, and reading (1c), where John greeted both Paul and Ben. We will follow linguistic convention and call the first form of ellipsis, where verb and grammatical object are elided 'gapping' (1b) and the second one 'conjunction reduction', or 'nongapping' (1c).

Gapping and Prosody

It has been noted in the literature that it can be very hard for listeners sometimes to arrive at the gapping interpretation of an ambiguous sentence; indeed, the nongapping version seems to be highly preferred. For instance, Carlson (2001) showed in a written questionnaire study that in sentences very similar to (1a), gapping interpretations are chosen only 4% of the time. Unfortunately, she did not include these sentences in a replication experiment where stimuli were presented auditorily, and prosody was manipulated to either bias towards the gapping or towards the nongapping interpretation. Bias was accomplished by making use of the fact that placement of pitch accent correlates strongly with the presence of new or contrastive information (e.g., Lambrecht, 1994).

Let us take a closer look at sentence (1a). In both the gapping and the nongapping interpretation there are two *pairs* of elements in each conjunct that are contrasted. In the gapping reading (e.g., 1b), *John* and *Ben* make up the first pair of contrasted elements, and *yesterday* and *today* constitute the second pair of elements that is contrasted. In the nongapping condition (e.g., 1c), however, the first pair of contrastive elements is different: it is formed by *Paul* and *Ben* (i.e., John sees *Paul* on one day, and John sees *Ben* on another). Thus, the gapping and the nongapping reading of an ambiguous sentence seem to have distinct prosodic realizations by which the listener can tell them apart. But does the listener use this information?

The answer is yes, according to Carlson's results, the listener does use these prosodic cues, but not to the extent that gapping can become the preferred interpretation; nongapping is always preferred, no matter how strong the prosodic bias might be. Gapping promoting prosody can raise the percentage of gapping responses to a maximum of 44%, still leaving a majority of nongapping responses. Carlson concludes that prosody is indeed an important factor in the processing of gapping structures, but that there must be another, stronger, factor at work to create this seemingly invincible preference for nongapping structures. What can this factor be?

Gapping and Context

Carlson concluded from her experiments that the major force resisting the gapping interpretation is the well-known *Minimal Attachment* principle: in case of ambiguity, choose

the simplest syntactic structure in terms of nodes in the respective syntactic trees. However, there are a number of reasons not to accept this conclusion. Besides the fact that different syntactic formalisms give rise to different predictions about the syntactic complexity of gapping and nongapping readings (under some linguistic theories there *is* no complexity difference), and the fact that choosing the simplest structure implies that all possible structures are built and compared, which seems to be computationally challenging at the very least, there is a strong competing explanation, namely one involving the principle of *Minimal Topic Structure*.

For instance, Hoeks, Vonk & Schriefers (2002) provided strong off-line and on-line evidence for their view that readers, and listeners alike, prefer to have one and only one topic in any given utterance (hence minimal topic structure), *unless* contextual or prosodic cues suggest there is more than one. A topic can be roughly defined as the thing the utterance provides information about (which most of the time is the grammatical subject of a sentence). This preference for a single topic is predicted to lead to processing difficulty when a sentence has not one but two topics, as is the case in Sentence Coordinations (e.g., *Nathan* [topic1] *helped Wilma and Tessa* [topic2] *laughed*). Presenting these sentences in a context promoting two topics, instead of the usual single topic, effectively eliminated this processing difficulty. This suggests that S-coordinations are difficult not because they are possibly syntactically more complex, but because they are pragmatically more complex in terms of topic structure.

The importance of context for the processing of gapping sentences has been signaled as early as 1976 by Kuno. However, the principles that Kuno identified as important were based on intuitive judgment and were not, at the time, tested empirically. More recently, Keller (2001) conducted two off-line acceptability experiments to investigate whether the right context could indeed increase the acceptability of gapping sentences. He found that *unambiguous* gapping sentences (such as, e.g., *She accompanied the boy to school and he to university*) could be made as acceptable as their nongapping counterparts by using a suitable context (e.g., *Where did Hanna and Michael accompany the boy to?*). Keller's experiments do not answer the question of whether gapping has indeed become the preferred structure in *ambiguous* structures, nor can they say anything (or only very implicitly so) about the factor prosody, as only written stimuli were used. In the present two experiments we do want to find out under what circumstances, if any, gapping can become the structure of choice for the listener, by manipulating both context and prosody of ambiguous gapping sentences. Doing so will enable us to gauge the strength of each of these factors that seem crucial to the processing of gapping structures.

Experiment 1

We conducted two experiments using a speeded auditory decision task, in which participants had to indicate as fast as possible what reading (i.e., gapping or nongapping) first came to mind after hearing the ambiguous target sentence.

The first experiment used sentences such as (1a), which in Carlson's questionnaire study received a gapping reading only 4% of the time. If we can achieve a gapping percentage of over 50% in this specific set of sentences, then we may really have identified the right prosodic and pragmatic factors. The second experiment used the reverse logic, in that context and prosody were put in place to increase the number of *nongapping* interpretations, even when nongapping meant semantically implausible (which constitutes a really strong constraint against any kind of interpretation!). But let us first look at the processing of plausible sentences.

Method

Participants Thirty-two native speakers of Dutch were paid for participating in this experiment (27 female; mean age 21 years, age range 17-25). This imbalance in gender should not cause problems, because recent research has shown that, though there may be gender differences in the perception of *emotional* prosody, the sexes seem to be equally proficient when it comes to understanding *linguistic* prosody, which is of concern here (Raithel & Hielscher-Fastabend, 2004).

Materials & Design For this experiment, 32 sets of mini-dialogues were constructed, each set consisting of four versions of a given dialogue. Four experimental lists were constructed with 8 experimental dialogues per condition, and no list containing more than one version of a given item. Added to these 32 dialogues were another 32 dialogues from a related experiment, which served as fillers for Experiment 1, and which will be discussed as Experiment 2 (see below).

The order in which experimental and filler items appeared was determined semi-randomly and was the same for each list. Each list was presented to an equal number of participants and each participant saw only one list. The experimental items for the first experiment appeared in four versions as exemplified below (2a-d). Note that English translations are given of the original Dutch stimuli (target sentences are structurally identical between languages). Please note also that CAPITALS indicate the presence of a pitch accent.

2a. **(GG) Gapping Context, Gapping Prosody**
Context: Wilma has bought a new house. With what did Nathan and Tessa help her ?
Target: NATHAN helped Wilma with PAINTING and TESSA with WALLPAPERING.
Proposition (Gapping): Tessa helped Wilma.

2b. **(GN) Gapping Context, Nongapping Prosody**
Context: Wilma has bought a new house. With what did Nathan and Tessa help her ?
Target: Nathan helped WILMA with PAINTING and TESSA with WALLPAPERING.
Proposition (Gapping): Tessa helped Wilma.

2c. **(NG) Nongapping Context, Gapping Prosody**

Context: Wilma and Tessa have each bought a new house. With what did Nathan help them ?
Target: NATHAN helped Wilma with PAINTING and TESSA with WALLPAPERING.
Proposition (Gapping): Tessa helped Wilma.

2d. **(NN) Nongapping Context, Nongapping Prosody**
Context: Wilma and Tessa have each bought a new house. With what did Nathan help them ?
Target: Nathan helped WILMA with PAINTING and TESSA with WALLPAPERING.
Proposition (Gapping): Tessa helped Wilma.

These dialogues were created by crossing the factors Context and Prosody.

Context The contexts employed here served three distinct purposes. First, they were designed to promote a certain topic-structure (either one single topic, promoting nongapping, or two contrastive topics, promoting gapping). Secondly, the contexts provide background knowledge, sometimes also called 'given' or 'old' information. For instance, the fact that Wilma was helped is made clear in the context sequence of (2a) and is therefore old information in the subsequent target sentence. An utterance does not consist merely of 'old' information, of course, but crucially instantiates a link between what is known and what is new. This new information expressed by an utterance is called the 'focus'. Often, the focus of a sentence coincides with its grammatical object, but other grammatical elements can receive focus too. For example, in a sentence with a nongapping reading such as (1c), not only the pair of direct objects (e.g., *Paul* and *Ben*) stand in a contrastive focus relation, but also the pair of adjunctive phrases (e.g., *yesterday* and *today*). So the third purpose of the context was to set up contrastive focus relations. Note that the pairs that are in focus are the same pairs that receive pitch accent (see *Gapping and Prosody* section above; see also *Prosody* section below).

Prosody Prosody is a multi-dimensional construct, and it would be very interesting to see what aspects of it (e.g., energy variations, duration of pauses etc.) play a role in the processing of gapping structures. However, for our present experiment we will focus on the role of pitch contours, which have been shown to crucially contribute to sentence comprehension. Figure 1 shows the pitch contours of the example sentence in both prosodic realizations. What distinguishes the two prosodies is the pattern of pitch accents in the first conjunct. In the prosody promoting gapping (see Figure 1, upper panel), the name of the direct object, Wilma, is de-accented because it is background knowledge, which is required if this information is to be left out of the second conjunct. The name of the subject, *Nathan*, does receive an accent, because it is part of the contrastive topic construction involving both *Nathan* and *Tessa*. Each of them is expected to perform some action. In the nongapping condition (see Figure 1, lower panel), the pattern of pitch accents in the first conjunct is completely reversed. The name of the grammatical subject,

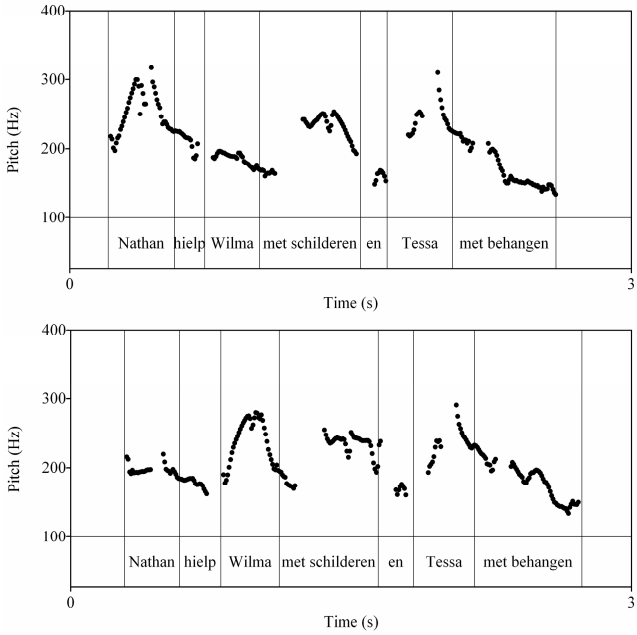

Figure 1. Pitch contours of an example target sentence. Upper panel: prosody promoting gapping; Lower panel: prosody promoting nongapping.

Nathan, is now de-accented, as it is the single topic of the sentence, and thus does not receive pitch accent (e.g., Lambrecht, 1994). In this case there are no contrastive topics. What is contrasted here is the direct object in the first conjunct, *Wilma*, which is set against the purported grammatical object in the second conjunct, *Tessa*. Because of this contrastive (focus-) relation, both receive pitch accent.

Two of the combinations (i.e., 2b and 2c) are in effect *infelicitous*, because they contain a mismatch between the expectations set up by the context and the actual prosodic realization in the target sentence. Comparing the results of these two specific conditions will show which one of the two factors is the most influential with respect to interpretation: the pragmatic context or the prosody of the target sentence.

All propositions in Experiment 1 represented the gapping reading of the ambiguous target sentence. The propositions belonging to the filler items, however, all stated the nongapping reading of the ambiguous target sentence. Thus, participants were presented with an equal number of gapping and nongapping structures (i.e., 32 of each). A practice session consisting of 16 dialogues preceded the actual experiment.

Procedure Participants were seated behind a computer screen in a sound-proof cabin. The dialogues were presented to them auditorily via two speakers. Each dialogue was preceded by a range of three asterisks appearing in the center of the screen ("***"), which indicated the start of a new item. After 1060 ms, the context sentence, spoken by a male speaker, was played, followed by the target sentence, spoken by a female speaker. The proposition (male speaker again) was played subsequently, together with a visual

presentation of three question marks ("???"), indicating that the participants should make a response.

Participants were instructed to indicate whether the proposition corresponded with the statement made by the *female speaker* (i.e., the target sentence), even if they thought the proposition sounded a little odd sometimes (this part was included with Experiment 2 in mind, see below). They could use the right SHIFT key on a keyboard for "YES" and the left SHIFT key for "NO". Participants were encouraged to respond as fast as they could and to follow their first impression; it was stressed that there were no 'correct' or 'incorrect' answers.

Results

Analysis Response times were calculated separately for "YES" responses and "NO" responses. In Experiment 1, a "YES" response is always a choice for the gapping reading; "NO" responses are choices for nongapping. The proportion of "YES" responses was also determined for each participant and each item. For all dependent measures, two analyses were performed: an *F1*-ANOVA on participant means for each condition and an *F2*-ANOVA on item means. The factors Context (promoting gapping vs. promoting nongapping) and Prosody (promoting gapping vs. promoting nongapping) were treated as within-participants and within-items factors. In addition, a planned comparison was performed between the two conditions where context and prosody biased towards different interpretations (see 2b and 2c). Mean response times (in ms) and mean proportions gapping response are presented graphically in Figure 2.

Proportion Gapping Responses The main effects of Context and of Prosody were highly significant on both participant- and item-analyses (all *p*-values < .0001). These results were qualified, however, by a significant two-way interaction between Context and Prosody ($F1(1,31)=8.20$, $p<.01$; $F2(1,31)=8.62$, $p<.01$). This interaction suggested that the effect of prosody favoring gapping was stronger when the context favored gapping (i.e., an effect of 25%), than when the context did not (i.e., an effect of 8%). Post-hoc tests revealed that both these effects were significant ($p<.05$). The planned comparison showed that the condition where the context biased towards gapping responses and the prosody did not (i.e., "GN") produced significantly more gapping responses than in the condition where the context did not favor gapping, but the prosody did (i.e., "NG"). The highest proportion of gapping responses, 61%, was found in the "GG" condition, where both context and prosody promoted gapping.

Response Times "YES" Responses Few participants gave a "YES" response in every condition (with fewest "YES" responses in the "NN" condition, see upper panel of Fig. 2), which led to a considerable number of empty cells in the Repeated Measures ANOVA. The number of participants upon which to base this analysis became so small (and with that, the statistical power of the test) that we decided to perform a Univariate ANOVA (which can be seen as stricter than the repeated measures ANOVA, as it tests against a larger error variance), with Context and Prosody as fixed

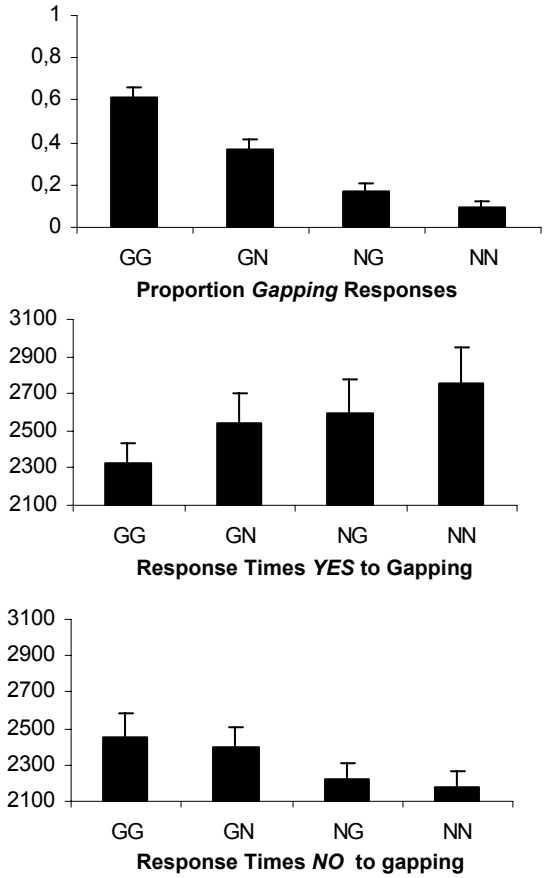

Figure 2. "GG"=Gapping context, Gapping prosody; "GN"=Gapping context, Nongapping prosody; "NG"=Nongapping context, Gapping prosody; "NN"=Nongapping context, Nongapping prosody.

between-group factors. The main effect of Context was significant in the analysis by items, though not by participants ($F1(1,83)=3.34$, $p=.13$; $F2(1,98)=10.18$, $p<.01$). This main effect reflected the shorter response times for accepting the gapping interpretation of the target sentence when the context promoted gapping (2431 ms; SE=96) than when it did not (2677 ms; SE=128). No main effect of Prosody was found, nor was there a significant interaction (all *p*-values > .20).

Response Times "NO" Responses Though Repeated Measures ANOVA's were possible as there were only few empty cells for the "NO" responses, we will only report the results for the Univariate ANOVA, for ease of comparison with the previous section (results of the Univariate and Repeated Measures analyses were identical). The main effect of Context was significant by participants and by items ($F1(1,120)=5.10$, $p<.05$; $F2(1,124)=25.63$, $p<.001$). Thus, when the context favored gapping , it took participants significantly longer to reject the gapping interpretation (2427 ms; SE=73) than when the context did not (2199 ms; SE=70). No main effect of Prosody was found, nor was there a significant interaction (all *F*-values < 1).

Discussion

In this experiment we wanted to find out whether it was possible to make the gapping reading the preferred interpretation of an ambiguous structure containing ellipsis. If we could, we should see an increase in percentage from about 4% in isolation (see Carlson, 2001) to over 50% in the present experiment. Indeed, we saw that given the right context and the right prosody, participants chose gapping over 60% of the time (which is significantly different from 50%; $p<.05$). Importantly, the pattern of reaction times convincingly shows that there is no sign of any speed-accuracy trade-off; on the contrary, there seems to be a strong linear relationship between proportion of gapping responses given, the ease and speed with which they are given, and the difficulty of rejecting the gapping response.

These results suggest that, even if structural considerations were to play a role in resolving ellipsis, they are subordinate to the combination of pragmatic and prosodic information. However, there must be some reason why the 61% that was found here is not 90% or more, which would be expected if context and prosody would have really won hands down. This could of course be due to matters of methodology: we used a task that required participants to consciously access a memory trace of (part of) a dialogue they had just heard. This procedure may give rise to considerable noise in the decision process, and thus to a less than perfect end-result. On the other hand, there may also be a theoretical factor at work which we have not addressed here. Future investigations must make this clear.

In addition, we wanted to know which one of the two factors, context or prosody, was actually guiding the process; in other words, which one was most important for choosing gapping? Our results are very clear in showing that prosody and context both have an *independent* contribution to the processing of gapping structures. For instance, the effect of prosody in promoting gapping responses is significant in both the gapping and the nongapping context. However, our results also indicate that in the case of context and prosody pointing in different directions, context wins. That is, there are significantly more gapping responses when context favors gapping and prosody does not ("GN"), than when it is prosody that favors gapping, and the context does not ("NG"). Furthermore, the analyses on response times were very clear in that they showed only a significant main effect of Context, but not of Prosody. Thus, pragmatic context outweighs prosody in the processing of sentences with ellipsis.

Experiment 2

The clear conclusion from Experiment 1 is put to a stringent test in Experiment 2. As was indicated earlier, this experiment uses a different logic than Experiment 1. In Experiment 2, the gapping response is the preferred response altogether, presumably only because choosing the normally prevailing nongapping reading results in semantic anomaly. Consider, for instance, dialogue (3a). If listeners choose the default nongapping reading, they will end up with a representation of the sentence reflecting that "some plasterer filled some painter", which of course does not correspond to the usual state of affairs in the world. So what we are interested to see in this experiment is whether the context and the prosody manipulations promoting *nongapping* responses are as strong as when they encourage gapping responses. They must perhaps even be stronger, because they have to take it up against semantic implausibility, which is indeed very likely to be a strong determiner of sentence interpretation in case of ambiguity.

Method

Participants The same thirty-two native speakers of Dutch who took part in Experiment 1.

Materials & Design The dialogues for this experiment appeared in four versions, see examples below (3a-d).

3a. **(GG) Gapping Context, Gapping Prosody**
Context: The wall was full of holes. What did the plasterer and the painter do?
Target: The PLASTERER filled the wall with a SPATULA and the PAINTER with a FILLING KNIFE.
Proposition (Nongapping): The plasterer filled the painter.

3b. **(GN) Gapping Context, Nongapping Prosody**
Context: The wall was full of holes. What did the plasterer and the painter do?
Target: The plasterer filled the WALL with a SPATULA and the PAINTER with a FILLING KNIFE.
Proposition (Nongapping): The plasterer filled the painter.

3c. **(NG) Nongapping Context, Gapping Prosody**
Context: The wall was full of holes, said the painter. What did the plasterer do?
Target: The PLASTERER filled the wall with a SPATULA and the PAINTER with a FILLING KNIFE.
Proposition (Nongapping): The plasterer filled the painter.

3d. **(NN) Nongapping Context, Nongapping Prosody**
Context: The wall was full of holes, said the painter. What did the plasterer do?
Target: The plasterer filled the WALL with a SPATULA and the PAINTER with a FILLING KNIFE.
Proposition (Nongapping): The plasterer filled the painter.

Results

See Figure 3 (next page) for a graphical presentation of mean proportions nongapping response and response times.

Proportion Nongapping Responses The main effects of Context and of Prosody were significant on both participant- and item-analysis (Context: $F1(1,31)=5.69$, $p<.05$; $F2(1,31)=22.76$, $p<.001$; Prosody: ($F1(1,31)=13.48$, $p<.01$; $F2(1,31)=13.26$, $p<.01$)). Participants accepted the implausible nongapping reading significantly more often (30% of the time) when the context indeed promoted nongapping than when it promoted the gapping reading (21% of the time). At the same time, there was an effect of prosody: the nongapping interpretation was chosen significantly more often when the target sentence was pronounced with a nongapping prosody (31% of the time) as

compared to when it had a gapping prosody (20% of the time).

The interaction of Context and Prosody was marginally significant by participants, and not significant by items ($F1(1,31)=2.82$, p=.10; $F2(1,31)<1$). This very weak interaction may be held to suggest a trend for a somewhat stronger effect of prosody in the nongapping context (i.e., a 13% increase in nongapping) than in the gapping context (i.e., a 9% increase). The planned comparison did not reveal a significant difference between "GN" and "NG".

Response Times "YES" and "NO"-Responses Statistical analysis revealed only one effect that approached significance, which was the main effect of Context in response times for "YES" responses (Univariate analysis). This main effect was marginally significant by items, but did not reach significance in the analysis by participants ($F1(1,60)=1.97$, p=.17; $F2(1,108)=2.84$, p=.095). The effect suggested a trend for participants to be faster to say "YES" to nongapping when the context favored nongapping, than when the context favored gapping (a difference of 227 ms). No other effects or planned comparisons were significant.

Discussion

This second experiment showed that even in the face of a very strong adversary such as semantic implausibility did pragmatic context and prosody have strong and independent effects on the interpretation of sentences with ellipsis. Indeed, if context and prosody teamed up to favor nongapping, it was chosen almost 40% of the time (which is barely different from 50%, p=.046)! This is really an accomplishment, given the enormous pressure not to choose an implausible interpretation. There was no significant difference between the two conditions with conflicting factors. Nevertheless, the evidence from the response times does suggest that Context was more important in the process of choosing the nongapping option than Prosody. This leads us to conclude that, again, pragmatic context is stronger than prosody in the processing of sentences with ellipsis.

Conclusion

These two experiments made very clear that nonstructural factors such as context and prosody are predominant in the processing of gapping structures. Experiment 1 showed that the right context and the right prosody can make gapping the *preferred* interpretation (gapping is chosen over 60% of the time). Experiment 2 showed that the right combination of pragmatic context and prosody can encourage listeners to choose the nongapping reading almost 40% of the time, even if this reading is highly implausible! Both experiments indicated that context and prosody have strong and *independent* effects on processing gapping. Context appears to be the stronger factor though: very clearly so in Experiment 1, admittedly rather less convincingly in Experiment 2. These findings establish pragmatic context and prosody as key factors in human sentence processing.

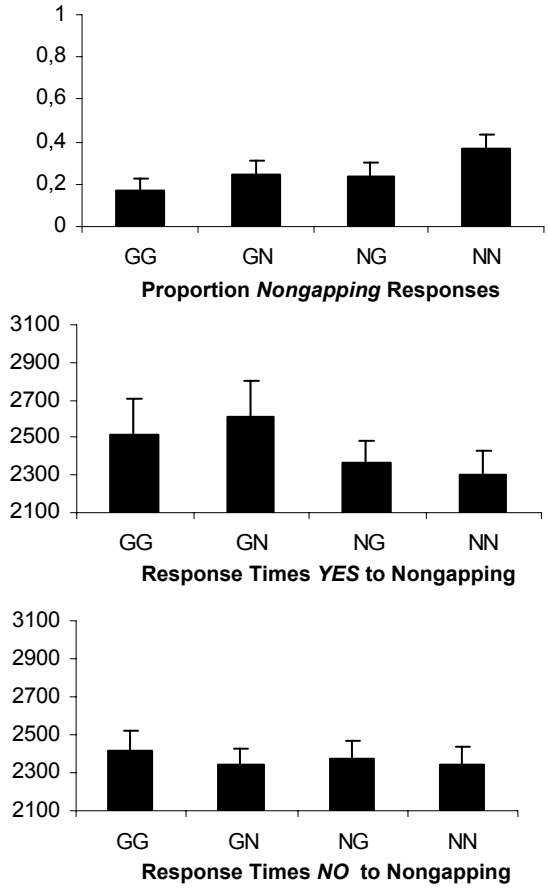

Figure 3. "GG"=Gapping context, Gapping prosody; "GN"=Gapping context, Nongapping prosody; "NG"=Nongapping context, Gapping prosody; "NN"=Nongapping context, Nongapping prosody.

Acknowledgements

This research was funded by grant # 015.001.103 from the Netherlands Organization for Scientific Research, awarded to Petra Hendriks. We are grateful to Erik-Jan Smits and Maartje Schreuder for lending their voices to our dialogues.

References

Carlson, K. (2001). The effects of parallellism and prosody in the processing of gapping structures. *Language and Speech, 44,* 1-26.

Raithel, V., & Hielscher-Fastabend, M. (2004). Emotional and linguistic perception of prosody. *Folia Phoniatrica et Logopedica, 56,* 7-13.

Hoeks, J.C.J., Vonk, W., & Schriefers, H. (2002). Processing coordinated structures in context. *Journal of Memory and Language, 46,* 99-119.

Keller, F. (2001). Experimental evidence for constraint competition in gapping constructions. In: G. Müller & W. Sternefeld (Eds.), *Competition in syntax* (pp. 211-248).

Kuno, S. (1976). Gapping. A functional analysis. *Linguistic Inquiry, 7,* 300-318.

Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents.* Cambridge, MA: University Press.