Prosody, Context, and Thematic Fit meet "Gapping": The Interaction of Multiple Constraints in Spoken Sentence Comprehension

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.

Gisela Redeker (g.redeker@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 prosody, context and thematic fit on off-line and on-line processing of sentences like *John greeted Paul yesterday and Ben today*. Such sentences are ambiguous between the so-called 'nongapping' reading, where *John greeted Ben*, and the highly 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. In addition, they respond to a visual probe that is presented during the presentation of the ambiguous target. The results show that context and prosody have independent and strong effects on both on-line processing and off-line interpretation, even if they are countered by thematic fit information.

Keywords: Psycholinguistics; Cross-Modal Gender Decision; Prosody; Pragmatics; Thematic Fit; Gapping.

Introduction

The language faculty is one of the most complex cognitive systems that we know of. This complexity arises from the stunning number of sources of information - that also are very different in nature - that have to be dealt with. Language production and comprehension require the integration of information about structure (syntax), about meaning and meaning combination (semantics), about the order (and sometimes morphology) of sentence elements in a sentence (information structure), about how the sentence fits into the linguistic and nonlinguistic context (pragmatics), not to mention factors such as prosody and visual scene information. Information coming from all of these sources has to be analyzed and integrated in order to succesfully utter a sentence or comprehend a message.

The linguistic phenomenon of gapping is - par excellence - an instance where many of these constraints are involved. Gapping is actually a form of ellipsis, where words or phrases are left out of an utterance without changing its formal meaning. Most of the time, ellipsis is associated with ease of processing, but gapping appears to be a special case in this respect. Consider, for instance, 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).

Why is Gapping so Hard?

In earlier studies it has been noted that it can be very hard for listeners 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 ambiguous sentences very similar to (1a), gapping interpretations are chosen only 4% of the time. Manipulation of the semantic parallelism between the entities in the sentence did lead to an increase in the number of gapping interpretations in sentences such as "John visited the office yesterday and Ben today", but only to 40%. Carlson (2001) suggested that the major force resisting gapping was the well-known *Minimal Attachment Principle*: choose the simplest syntactic structure in terms of nodes of the respective syntactic trees (Frazier, 1987). Because *non*gapping is syntactically less complex (i.e., its syntactic description contains fewer nodes), it should for that reason be the preferred option in case of ambiguity. Gapping only becomes the option of choice when thematic fit information is brought into play, as in "John baked cookies for his parents and Ben for his grandparents", where the nongapping interpretation is practically excluded. Here, 81% of instances received a gapping response, because the nongapping interpretation (i.e., John baked Ben) is highly implausible (Carlson, 2001).

Gapping and Prosody

In a follow-up experiment using auditory stimuli, Carlson (2001) manipulated the prosody of the ambiguous sentences, 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, as in "*JOHN greeted Paul YESTERDAY and BEN ~~greeted Paul~~ TODAY*" (capitalization here indicates the presence of pitch accent). 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), as in "*John greeted PAUL YESTERDAY and ~~John greeted~~ BEN TODAY*". Here, *John* is clearly de-accented. 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 raised the percentage of gapping responses to 44%, still leaving a majority of nongapping responses. Carlson concludes that prosody is indeed an important factor in the processing of gapping structures, but the very strong preference for nongapping structures is created by a structural factor, namely syntactic simplicity.

Gapping and Context

Keller (2001) conducted off-line acceptability studies to investigate whether, instead, contextual factors may be responsible for the unpreferredness of gapping, and whether providing the right context could indeed increase the acceptability of gapping sentences (see also Kuno, 1976). He found that gapping sentences made *unambiguous* by case marking (such as, e.g., *She accompanied the boy to school and he to university*) were significantly less acceptable than nongapping sentences (such as, e.g., *She accompanied the boy to school and the girl to university*, which, by the way, is still ambiguous). Only by using a specific context (e.g., *Where did Hanna and Michael accompany the boy to?*) could these unambiguous sentences be made as acceptable as their nongapping counterparts. Why would context have this effect?

Question contexts such as those used by Keller set up a complex expectation regarding the *information structure* of the answer, making the gapping interpretation of the subsequent sentence acceptable and perhaps even natural. This complex expectation concerns at least three different aspects of information structure: First of all, the context for a large part determines what will be the most likely *topic* of the following sentence, where 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). Hoeks, Vonk & Schriefers (2002) provide 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 (as in nongapping), unless contextual or prosodic cues suggest there is more than one.

This preference for a single topic is predicted to lead to processing difficulty when a sentence has not one but two topics, as is, for instance, the case in Sentence Coordinations (e.g., *John* [topic1] *greeted Paul and Ben* [topic2] *laughed*). Presenting these sentences in a context promoting two topics (e.g,. *"What did John and Ben do?"*), instead of the usual single topic, effectively eliminated this processing difficulty. This suggests that S-coordinations are difficult not because they are possibly more complex syntactically, but because they are more complex in terms of topic structure. This is an important finding, because gapping, much like Sentence Coordination, involves having two contrastive topics (e.g., *John* [topic1] *greeted Paul yesterday and Ben* [topic2] *~~greeted Paul~~ today*). If the preceding context prepares the way for two topics, processing difficulty will decrease.

The second important feature of the context question is that it determines what will be the *focus* of the anwer, with focus being the new information that is provided by the anwer. In the gapping interpretation of the ambiguous sentence (1a), there are contrastive foci set up by the question word 'Where', suggesting that the answer will contain (prepositional) phrases indicating locations, one location for John's action, and one location for Ben's action.

Finally, the question context makes unambiguously clear that the fact that "Paul was greeted" is given information. This is important, because only information that is given, can be 'gapped', i.e., left out of the sentence (Kuno, 1976).

Thematic Fit, Context, and Prosody

Recently, Hoeks, Hendriks, and Zijlstra (2006) investigated the interaction of the three factors that seem most critical to processing gapping, namely thematic fit, pragmatic context, and prosody. They used a so-called 'auditory decision' paradigm, in which participants are presented with spoken mini-dialogues: one speaker asks a question, and another speaker answers that question. Pragmatic expectation was manipulated via the question, and prosodic structure of the answer was varied to either conform to the gapping or to the nongapping interpretation. Immediately after the offset of the answer, participants had to react to a spoken statement that was either consistent with the gapping or with the nongapping interpretation of the answer.

With respect to sentences without thematic fit manipulation such as (1a), Hoeks et al. found that, with the right context and the right prosody, gapping could indeed the preferred interpretation: over 60% of all cases were interpreted as gapping. Responses in this condition were also faster than in the other conditions, suggesting that the gapping option was wholeheartedly accepted by the listener. Recall that until this experiment no manipulation produced more than 44% gapping responses (Carlson, 2001). In a second experiment using sentences that were strongly biased towards gapping by means of thematic fit of the ambiguous NP with the matrix verb (e.g., *John baked cookies for his parents and Ben for his grandparents*), context and prosody still significantly influenced processing: gapping responses decreased from 83% to 63% when context and prosody biased against gapping, which is quite an accomplishment

given the strong pressure in general to choose plausible interpretations. This means that context and prosody were able to cause an increase of about 20% in choosing the implausible nongapping interpretation (i.e., where John baked Ben) of an ambiguous sentence.

Nevertheless, despite the evidence for the strong (combined) effects of context and prosody, and despite the fact that it was possible to make gapping the structure of choice in ambiguous sentences, these results were obtained via an off-line task. This makes it difficult to draw hard and fast conclusions about what is going on during the actual on-line processing of the ambiguous sentence. In other words, because the task measured off-line interpretation it does not permit us to choose between two options: 1) the preference for nongapping did not go away, and did lead to processing difficulty during on-line processing, but due to the available contextual and prosodic information, reanalysis was completed before the end of the sentence, and 2) the preference for nongapping was overruled by context and prosody from the start, so there was no processing difficulty whatsoever. To address this issue, two experiments were designed that basically used the same materials as the Hoeks et al. study, but this time an extra task was added in order to probe on-line processing: cross-modal gender decision.

Experiment 1

In both experiments that are reported here, we used 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 (for more details, see Hoeks et al., 2006). However, we also employed an additional task involving the presentation of a visual probe while participants listened to the target sentence, to find out whether there are differences in on-line processing between gapping and nongapping. Two kinds of visual probes, in both cases proper nouns, were presented near the end of the sentence to estimate the processing difficulty associated with gapping, and participants had to decide both quickly and accurately whether a probe was a girl's name or a boy's name, hence the term 'gender decision'. How does it work?

Suppose we have a sentence with gapping interpretation, such as "*Nathan helped Wilma with painting and Tessa* ~~helped Wilma~~ ***with*** *wallpapering*". The first kind of probe is identical to the proper noun contained in the gapped constituent ('helped Wilma'), which in this example is the name 'Wilma'. The idea behind this is that when listeners compute the gapping interpretation, which we assume will happen in the vicinity of the second preposition (i.e., the 'with' of 'with wallpapering'), the concept denoted by 'Wilma' will be activated in some way or the other, or will at least be consistent with the representation that is being built right then. In contrast, when listeners process the nongapping interpretation, "*Nathan helped Wilma with painting and* ~~Nathan helped~~ *Tessa* ***with*** *wallpapering*", there is no reason to assume that the concept associated with 'Wilma' is activated at the second preposition, at least not to the same degree as in the gapping interpretation. So with respect to this *identity* probe, we expect faster recognition in conditions with a predominant gapping response as compared to conditions where nongapping is the prepotent response.

The other type of probe, the *control* probe, bears some superficial phonological resemblance with the identity probe, but is not identical to any of the names used in the experiment; in the example at hand this control probe would be something like 'William'. The recognition of the control probe does not depend on the level of activation of the concept belonging to 'Wilma', we assume instead that recognizing this control probe is mainly sensitive to the amount of processing difficulty occurring in the sentence where the probe is placed. In other words, control probes measure processing load.

To try and avoid floor effects for recognizing the identity probes, the probes were visually 'degraded' by using a different font size and font style for every character of the name (also mixing capitals and normal letters). Furthermore, we reduced visual contrast by presenting the probes in lightgray on a white screen. The rationale is that by making the process of recognition longer and more difficult, effects of priming should become more visible. The same reasoning underlies our choice for control probes that are to some extent phonologically similar to the identity probes: having partially similar probes should encourage our participants to make a decision only when the probe is fully identified, and not to give a reaction based on partial identification (for instance, to infer that the probe is 'Wilma', when they've only seen 'Wil'). To avoid memory related strategies we did not ask participants to say whether or not they think the probe occurred in the sentence they just heard, but merely to decide whether the name was typically a boy's or a girl's name.

Method

Participants Thirty-six native speakers of Dutch were paid for participating in this experiment (24 female; mean age 21 years, age range 19-27). The imbalance in gender should not cause problems with respect to the processing of prosody, because 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, 48 sets of mini-dialogues were used, each set consisting of three versions of a given dialogue. Since we had three gapping conditions (see 2a-c below) and two probe conditions (identical vs. control), six experimental lists were constructed, to make Condition and Probe Type within-participants variables. Each of the six lists had 8 experimental dialogues per condition, and no list contained more than one version of a given item. Added to these 48 dialogues were another 32 dialogues, 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 three versions as exemplified below (2a-c). 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. Probes are signaled using brackets: < *Wilma*>.

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 <*Wilma / Willem*> 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 <*Wilma / Willem*> WALLPAPERING.
Proposition (Gapping): Tessa helped Wilma.

2c. **(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 <*Wilma / Willem*> WALLPAPERING.
Proposition (Gapping): Tessa helped Wilma.

All propositions at the end of the dialogues in Experiment 1 represented the gapping reading of the ambiguous target sentence. The propositions belonging to Experiment 2, however, all stated the nongapping reading. As to the probes: control probes were chosen such that they matched the identical probes as closely as possible in terms of number of characters and Google frequency. The practice session that preceded the actual experiment consisted of three parts. In the first part the gender decision task was practiced without any auditory input, then the auditory decision taks with the statements was practiced without the gender decision task, and then the tasks were combined.

Procedure Participants were seated behind a computer screen in a sound-proof booth. The dialogues were presented to them aurally via two speakers. Each dialogue was preceded by a range of three asterisks appearing in the center of the screen ("***"). 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.

At the off-set of the second preposition in the ambiguous target sentence (e.g., the 'with' of the phrase 'with wallpapering' in 2a-c), either an identical or a control probe was presented which remained on screen until a decision was made. For this task, participants also used the SHIFT keys: right SHIFT for "boy's name" and LEFT shift key for "girl's name". To avoid confusion, the SHIFT keys were labeled, and A5-sized pieces of paper were placed to the right and to the left of the keyboard with "yes - boy's name" and "no - girl's name", respectively.

Results

Analysis Proportions "YES" to the *statements* were calculated for each participant and each item. In this experiment, a "YES" response is always a choice for the gapping reading; "NO" responses are choices for nongapping. In addition, response times to the *probes* were calculated, but only for correct responses (i.e., where the chosen gender matched the gender of the probe). For all dependent measures, we performed ANOVA's on participant and item means. The factors Condition (see 2a-c: "GG", "GN", and "NN") and Probe Type (identical vs. control) were treated as within-participants and within-items factors. Mean proportions gapping response are presented in Table 1, together with the data from the Hoeks et al. (2006) experiment, to show that the patterns of results are very similar. The gender decision times are presented in Figure 1.

Table 1: Mean Proportions of Gapping (SE)

Condition	Hoeks et al.	Exp 1
GapCont-GapPros	0.61 (.05)	0.62 (.05)
GapCont-NongapPros	0.37 (.05)	0.45 (.04)
NongapCont-NongapPros	0.09 (.03)	0.16 (.03)

Proportion Gapping Responses The main effect of Condition was highly significant on both participant- and item-analyses ($F1(2,60)=58.69$, $p<.001$; $F2(2,84)=168.61$, $p<.001$). All conditions differed significantly, and the highest proportion of gapping responses, 62%, was found in the "GapCont-GapPros" condition, where both context and prosody promoted gapping.

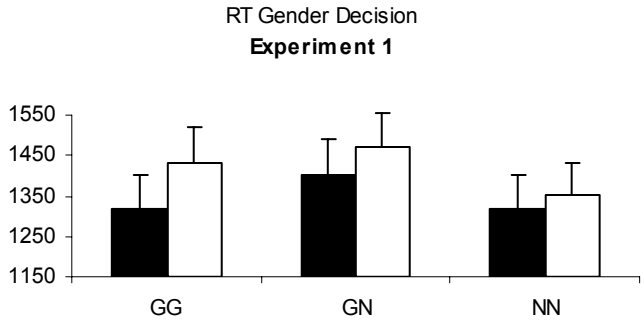

Figure 1. "GG"=Gapping context, Gapping prosody, etc. black bars = identical probe; white bars = control probe.

Gender Decision Times Participants were surprisingly good at the dual task: on average they were correct 96% of the time. Statistical analysis of the gender decision times produced a significant main effect of Probe Type ($F1(1,30)=11.02$, $p<.01$; $F2(1,42)=4.12$, $p<.05$), indicating that the decisions to the identical probes (1346 ms; SE=83) were faster than to the control probes (1419 ms; SE=80). We had expected this priming effect to be different for the different conditions, but there was no statistically reliable interaction of Condition and Probe Type (both *p*-values > .25). Instead, the main effect of Condition was significant ($F1(2,60)=6.67$, $p<.01$; $F2(2,84)=5.45$, $p<.01$), indicating that decision times to probes, whether or not they were identical, were slowest in the GN condition, but did not differ between GG and NN. This pattern was supported by post-hoc tests.

Discussion

First of all, the results of the responses to the statements constituted an almost perfect replication of the pattern of the off-line results found by Hoeks et al. (2006). When gapping context and gapping prosody conspired, participants chose the gapping interpretation for over 60% of the time, whereas when context and prosody biased towards nongapping, gapping clearly remained unpreferred, and was chosen only 16% of the time (with the mismatch condition somewhere in between). Taken together with the very few mistakes that were made in the gender decision task, this result shows that participants were able to handle this dual task paradigm extremely well.

With respect to on-line processing, we found a significant identity priming effect, but we did not find the expected difference in priming between the three conditions. Recall that we assumed the gapped elements (e.g., 'helped Wilma' in the example) would be (re-)activated during the processing of the gapping interpretation, which should have made it easier to process the identity probe (e.g., 'Wilma'). This null-effect could mean that, despite our efforts to circumvent it, we did get a floor-effect for the identity probes, meaning that although there may have been a difference in reactivation between the conditions, there is an upper limit as to the speed of gender decision times.

An alternative view is that there actually was no difference in reactivation between the conditions. If the elided material is not reactivated in its surface form, but only in some kind of interpreted form, it may not yield the differential priming effect we were looking for. If these assumptions are true, than the probes will measure processing load rather than activation. The results then become rather clear: computing the gapping interpretation is as easy as computing the highly preferred nongapping interpretation, if the right context and prosody are applied. The only processing difficulty is found in the mismatch condition, where context and prosody point to different directions.

Experiment 2

This second experiment uses a different logic than the first. To start with, because of the thematic fit manipulation, the gapping response is the preferred off-line response, 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. In addition, we present visual probes while participants listen to those target sentences, to look at what is happening during on-line processing of the ambiguous sentences where the preference for nongapping collides head on with the preferred thematic structure. In this experiment we are not looking for reactivation patterns (the materials contain no proper nouns); the probes are solely used to measure processing load.

Method

Participants & Materials & Design The participants of Experiment 1 also took part in this experiment. The materials are dialogues appearing in four versions (e.g., 3a-d). Probes are represented as follows: <*Walter*>.

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 <*Walter*> 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 <*Walter*> 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 <*Walter*> 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 <*Walter*> a FILLING KNIFE.
Proposition (Nongapping): The plasterer filled the painter.

Results

See Table 2 for mean proportions nongapping and Figure 2 for a graphical presentation of gender decision times.

Table 2: Mean Proportions of Nongapping (SE).

Condition	Hoeks et al.	Exp 2
GapCont-GapPros	0.17 (.06)	0.17 (.05)
GapCont-NongapPros	0.25 (.06)	0.26 (.06)
NongapCont-GapPros	0.24 (.06)	0.30 (.06)
NongapCont-NongapPros	0.37 (.06)	0.38 (.06)

Proportion Nongapping Responses The results of this experiment are virtually identical to those of Hoeks et al. (2006). Again, there were main effects of Context and of Prosody (Context: $F1(1,32)=12.94$, $p<.001$; $F2(1,28)=53.54$, $p<.001$; Prosody: ($F1(1,32)=9.62$, $p<.01$; $F2(1,28)=24.08$, $p<.001$)). Participants accepted the implausible nongapping reading significantly more often (34% of the time) when the context indeed promoted nongapping than when it promoted the gapping reading (22% of the time). The effect of prosody entailed that the nongapping interpretation was chosen significantly more often when the target sentence was pronounced with a nongapping prosody (32% of the time) as compared to when it had a gapping prosody (23% of the time). There was no interaction of Context and Prosody.

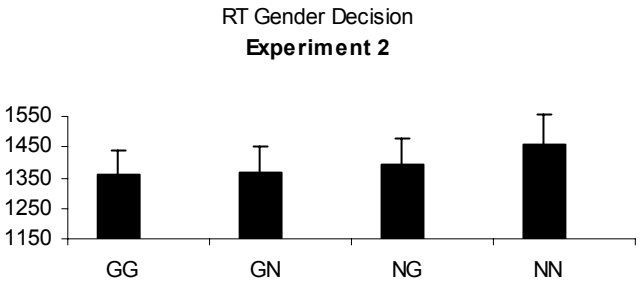

Figure 2. "GG"=Gapping context, Gapping Prosody etc.

Gender Decision Times As in Experiment 1, participants were really good at the task: on average 96% of the reponses were correct. Statistical analysis of gender decision times revealed that there was no effect of Prosody. There was a main effect of Context which was significant by participants and marginally significant by items ($F1(1,32)=4.66$, $p<.05$; $F2(1,28)=3.71$, $p=.06$) showing that participants are slower to respond to the probe in conditions where the context goes against gapping (1425 ms, SE=87) than when it promotes it (1363 ms, SE=80). This main effect was qualified by a trend towards an interaction (but only in the analysis by items) between Context and Prosody ($F1(1,32)=1,33$, $p=.26$; $F2(1,28)=3.88$, $p=.06$): the most difficult condition, where both context and prosody worked against the plausible gapping interpretation, showed the longest decision times; the other three conditions did not differ significantly.

Discussion

The data regarding the final interpretation of the ambiguous sentences (the off-line auditory decision data) are very similar to those reported by Hoeks et al. (2006). In line with these earlier results, the present experiment showed clear main effects of both Context and Prosody, indicating that both factors contribute significantly and independently to the interpretation of ambiguous sentences if strong thematic fit information is present. However, the gender decision data are rather surprising. As could have been expected, there is a significant disadvantage for the most difficult condition (i.e., "NN", as in (3d)), where both context and prosody argue against the gapping preference induced by thematic fit. However, quite unexpectedly, there is no evidence whatsoever for an *advantage* in the easiest condition (i.e., "GG" as in (3a)), where all three factors: thematic fit, context, and prosody argue for gapping. This may mean that there is only very little processing difficulty in all but the "NN" condition, which is solved extremely rapidly, before the probe is encountered. On the other hand, it may be the case that processing difficulty is present for each of the three 'faster' conditions, irrespective of the strength of the different factors. We are currently planning an experiment with structurally unambiguous sentences (using 'while' instead of 'and') to settle this issue.

Conclusion

The present experiments show that nonstructural factors such as context and prosody have strong effects on both on-line processing and off-line preferences of gapping sentences. In good thematic fit sentences, context and prosody seem to eliminate all processing difficulty associated with gapping, and make it the structure of choice. In case of poor thematic fit, context and prosody have a significant effect only when *both* go against the gapping bias induced by thematic fit. What is unclear at present is why the mismatch conditions are as easy (or as difficult) as the condition where all factors are fully in line with thematic fit. More research is clearly needed to fill this gap.

Acknowledgements

This research was funded by grant # 015.001.103 from the Netherlands Organization for Scientific Research, awarded to Petra Hendriks. We are very grateful to Louisa Zijlstra, Tim Bosselaar, Erik-Jan Smits, and Maartje Schreuder for their invaluable help with creating the stimuli and running the experiments.

References

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

Frazier, L. (1987). Sentence processing: A tutorial review. In: M. Coltheart (Ed.), Attention & Performance XII (pp. 601-681). Hove, UK: Lawrence Earlbaum.

Hoeks, J.C.J., Hendriks, P., & Zijlstra, L.J. (2006). The predominance of nonstructural factors in the processing of gapping sentences. In: R. Sun (Ed.), *Proceedings of CogSci 2006* (pp. 1511-1516).

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.* Cambridge, MA: University Press.

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