Gerrit Kentner (Potsdam), Caroline Féry (Potsdam) and Kai Alter (Newcastle upon Tyne)

Prosody in Speech Production and Perception: The Case of Right Node Raising in English

1 Introduction

Prosody, as an intrinsic feature of spoken language, provides information about both the underlying syntactic structure and the information structure of an utterance (e.g. Nespor & Vogel 1986, Beckman 1996, Ferreira 2002). With respect to syntactic and information structures, studies on prosody have mainly examined two prosodic features: the distribution of intonational phrase (IP) boundaries and the position and type of accents.

The distribution of IP boundaries is dependent on the underlying syntactic structure as well as the amount of phonetic material to be realised within the utterance (Ferreira 2002, Watson & Gibson 2004). Several studies have shown that listeners can use these boundaries as a signal to syntax in on-line sentence processing (eg. Warren et al. 1995, Pynte & Prieur 1996, cf. Cutler et al. 1997 for a review). These studies are in line with models that assign prosody a role as input structure for the parser (Marcus & Hindle 1990, Speer et al. 1996, Schafer 1997). Accordingly, the syntactic structure of a sentence is computed on the basis of its prosodic structure.

The syntax-prosody interplay is complicated by findings according to which speakers do not always reliably encode prosodic boundary cues in the signal (Albritton et al. 1996, Snedeker & Trueswell 2003). Therefore, it remains unclear whether prosody is necessary or even helpful for sentence parsing. Given these restrictions, it seems possible that syntactic parsing initially takes place without recourse to prosody. In this *syntax first* view on sentence processing, prosody merely serves as a filter, confirming or rejecting purely syntactic parsing results.

The distribution of pitch accents reflects the intonational highlighting of discourse relevant information and is thus determined by the information structural environment. In English, given an information structurally neutral environment, pitch accents are located at a (syntactically defined) default position in the right periphery of the sentence (nuclear stress, cf. Chomsky & Halle 1968). This is the case in sentences which are uttered “out of the blue” and that do not refer to any pre-established discourse context. This instance of “broad focus” (Ladd 1996) contrasts with two kinds of narrow foci (Toepel & Alter 2004). Narrow-new focus concerns constituents that are new in the discourse but expected as an answer to wh-questions. Contrastive focus, on the other hand, is an instance of narrow focus in which one element is contrasted to a set of possible alternatives that are established by the discourse. Contrastive focus thus does not convey new information.

Both instances of narrow focus are signalled with pitch accents, although the type of accent may be different from broad focus accent (Alter et al. 2001, Selkirk 2002). In narrow foci, pitch accents fall on the focused constituents. Accent on the syntactically defined default position can be missing (deaccentuation, cf. Ladd 1996). Other than in broad focus contexts, the focus domain cannot exceed narrowly focused elements. However, contrastive focus can be embedded within a broader focus domain (Féry & Samek-Lodovici 2006).

Regarding the processing of pitch accents, it has been shown that listeners parse sentences more easily when pitch accents fall on constituents representing new information as compared to pitch accentuation of given constituents (e.g. Terken & Nooteboom 1987).

Carlson (2001) has shown that listeners use parallel pitch accentuation to compute contrastive focus in gapping structures. She has compared listener`s reaction on ambiguous gapping-sentences like (1) in two conditions.

(1) Bill took chips to the party and Susan to the game.

Listeners tend to interpret *Susan* as the subject of the second conjunct when both *Bill* and *Susan* are pitch accented. In the case of pitch accentuation of *chips* and *Susan*, respectively, *Susan* is interpreted as the sentential object. These findings are especially interesting because they show that listeners use pitch accents not only to compute the information structure of the utterance but also to indirectly derive its syntactic structure.

As is the case with prosodic boundary cues, the reliability of pitch accents as anchor points for the human sentence processor is debatable. Hruska et al. (2001) show that, under certain circumstances, the parser is deaf towards pitch accents. In light of this finding, we cannot be sure about the status of pitch accents as input structure for the parser.

The present investigation is concerned with the processing of prosody in one type of elliptical construction, Right Node Raising (RNR). In general, ellipses are considered comparatively demanding for the sentence processing system, since information that is necessary for interpretation remains unexpressed at the surface. As opposed to comparable non-elliptical sentences which display the same linear word order but different prosodic renderings, these ellipses can be seen as a test bed for the evaluation of the role of prosody in sentence processing.

The next section introduces the RNR structure, emphasising both its prosodic and its intrinsic information structural properties. A production experiment was devised to specify the prosodic structure of RNR sentences. In a perception experiment, the role of prosody for the recovery of the missing material in the elliptical sentences is examined. Our analysis especially deals with the interaction of different prosodic parameters (accentuation and phrasing) with the underlying syntactic structure during sentence processing. The results may specify the role of prosody in theories of sentence processing.

2 Right Node Raising

RNR constructions[1] are instances of ellipsis in co-ordination structures. In these constructions, one element interpreted as part of both conjuncts, appears in only one conjunct and is omitted in the other. The target of elliptical deletion in RNRs is the element at the end of the first conjunct. Its identical counterpart appears at the corresponding site of the second conjunct, namely at the right periphery. In (2) an example of RNR is given. The target of ellipsis is represented by crossed-out words.

(2) Ian has lost ~~a camera~~ and Stephen has found a camera.

RNR sentences display a complex focus structure with contrastive focus embedded in a broad presentational focus (Féry & Samek-Lodovici 2006). The pre-elliptical element in both conjuncts is contrastively focused. We follow Selkirk (2002) in marking broad focus by ***foc*** and contrastive focus by **FOC**. In (3) the differently focused elements of (2) are represented:

(3) Ian has [[lost]$_{\text{FOC}}$ ~~a camera~~]$_{\textbf{\textit{foc}}}$ and Stephen has [[found]$_{\text{FOC}}$ a camera]$_{\textbf{\textit{foc}}}$

The focus structure of (3) contrasts with the focus structure of a non-elliptical sentence with the same basic word order (4). As opposed to (3), in (4) broad and contrastive foci coincide within the verb phrase (VP).

(4) Phil has [[dined]$_{\text{FOC}}$]$_{\textit{foc}}$ and Sarah has [[written a letter]$_{\text{FOC}}$]$_{\textit{foc}}$

According to its focus structure, the prosodic realisation of a RNR sentence differs from the prosody of comparable non-elliptical sentences. Contrastively

---

[1] We stick with the term RNR although we do not assume syntactic movement to apply in these structures. An alternative term is ‘backward deletion’ (Wilder 1997).

focused elements attract prominence. Thus, in (4) the VP-internal object (*letter*) bears the main stress. This prominence pattern corresponds to nuclear stress, according to which the rightmost element within the focus be accented. In (3), however, the verb (*found*) is most prominent within the focused VP, leaving the object to its right unaccented. The crucial difference lies in the contrastive focus, which, in RNR sentences, obviously overrides the nuclear stress rule (Féry & Samek-Lodovici 2006).

Selkirk (2002) has examined the prosodic structure of English RNR sentences. Her results show a preponderance of L+H* pitch accent on the contrasted elements in both conjuncts, followed by a low boundary tone. There are, however, two caveats with respect to Selkirk's results. Firstly, she collapses the data for the contrasted elements of both conjuncts. For example, she reports a pause after 35% of the contrasted elements, but it remains unclear to what extent the distribution of pauses differs in the two conjuncts. Secondly, the presence of a phrase boundary after the contrasted element in the second conjunct probably depends on the length of the elliptical target (cf. Hartmann 2000). Short targets tend to be integrated into the current IP whereas long targets will make up their own IP. Selkirk has not controlled for the length of the targets in her material.

2.1. Open Questions/ Hypotheses

As shown above, there is empirical evidence that, with respect to the second conjunct, RNR sentences differ prosodically from comparable non-elliptical sentences, reflecting their differences in focus structure. However, it is not clear yet, whether there is a prosodic difference already at the end of the first conjunct (the ellipsis site in RNR sentences). RNR sentences and their non-elliptical counterparts have different syntactic structures. This difference might reflect itself in phonetic structure, thereby possibly indicating the presence of an ellipsis. This hypothesis will be dealt with in a speech production experiment.

It is assumed that elliptical sentences in general are relatively difficult to parse, since material necessary for interpretation must be recovered non-locally. The question arises whether prosodic cues can help listeners to circumvent these difficulties. Another question is whether prosodic cues at the ellipsis site (i.e. at the end of the first conjunct) play a role in parsing.
A speech perception experiment was conducted to investigate these issues and to specify listeners' capacity to exploit prosodic cues to interpret elliptical sentences.

3 Speech Production

In order to describe the specific prosodic structure of RNR sentences, they were compared to non-elliptical sentences with the same linear word order. The sentences employed are ambiguous with respect to a non-elliptical and a RNR interpretation. They all obey the following constituent order:

[PropName1][Aux][VP1][and][PropName2][Aux][VP2][object- NP]

VP1 is the source of the ambiguity, licensing a transitive as well as an intransitive reading. All words used were controlled for frequency (>10 per million according to CELEX database) and for number of syllables and rhyme type (either disyllabic trochees or monosyllabic words).

The sentence types and their respective interpretation are exemplified in (5), with a paraphrase of the first conjunct for the non-elliptical reading in (5') and for the RNR reading in (5'').

(5) Nina is riding and Ian is fixing a bike

(5') Nina is riding on a horse

(5'') Nina is cycling

3.1 Speech Recordings

Thirty-seven ambiguous sentences were constructed according to the above pattern. The sentences were read aloud by a female speaker in an acoustically shielded room. The speaker was instructed to read each sentence in both the non-elliptical and the RNR condition, thereby establishing a broad focus interpretation in both sentence types (as if to utter each sentence as an answer to "what's happening?"). The material was digitised at 44.1 kHz, 16 bit sampling rate.

3.2 Acoustic Analyses

Acoustic features of both sentence types were examined at two critical points. Firstly, the prosodic structure of the second conjunct, namely the VP and the object NP1, was investigated. Secondly, prosodic features of VP1 and at the end of the first conjunct were analysed in detail.

With respect to the second conjunct, the length of both VP2 and object were measured and compared across conditions. Moreover, the difference of the pitch maxima on VP2 and NP was calculated and compared.

As to the first conjunct, the following four parameters were scrutinised:

- Length of VP1
- Pause after VP1
- Pitch maximum on VP1
- Pitch course at offset of VP1[2]

The respective values were subjected to statistical analysis.

3.3 Results

The results for the production experiment are summarised in table 1. The critical constituents of the second conjunct exhibit significantly different acoustic parameters in the two conditions.

			non-ellipt.	RNR	t- value	p- value
Conjunct 2	Length (VP2)	ms	332	358	3.97	<0.001*
	Length (Obj)	ms	474	459	2.24	0.03*
	F0max (VP2)- F0max (Obj)	Hz	12.92	27.32	4.84	<0.001*
Conjunct 1	Length (VP1)	ms	470	440	4.16	<0.001*
	Pause (incl. „and“)	ms	260	150	7.34	<0.001*
	F0max (VP1)	Hz	222	225	1.184	0.24
	F0min - F0off (VP1)	Hz	-14.24	-1.62	6.15	<0.001*

Table 1: Statistical analysis of prosodic features of non-elliptical and RNR- sentences

The contrastively focused VP2 in RNR sentences is lengthened as compared to VP2 in the non-elliptical condition, whereas the object in RNR sentences is shorter than its counterpart in non-elliptical sentences. Furthermore, analyses of the pitch pattern show that in the case of RNR, VP2 bears a relatively higher pitch accent with respect to the object than in non-elliptical sentences.

The examination of the first conjunct also yields a prosodic disparity across conditions. Although there is no significant difference with respect to the absolute pitch maximum on VP1, the phrasing structure reveals a striking contrast between the two conditions. The duration of VP1 is distinctly longer in the non-elliptical version, and so is the pause after VP1. In addition, the pitch course displays a F0-rise after the F0-minimum on VP1 in the non-

[2] The difference between the pitch minimum and the pitch at the offset of VP1 was calculated. This calculation yields information about the boundary tone. In case of a low boundary tone, the difference is near 0 (offset and minimum coincide), whereas a higher amount indicates a rise in pitch.

elliptical condition. All parameters investigated indicate a strong intonational phrase boundary for the non-elliptical sentences, marked by final lengthening with a high boundary tone and a pause.

In the RNR-condition, however, the boundary between the conjuncts is significantly weaker. Final lengthening on and pause after VP1 are less pronounced. F0 minimum and F0 at offset of VP1 mainly coincide, reflecting a low boundary tone between the conjuncts in this condition. Figure 1 and 2 show typical pitch tracks for the two conditions.

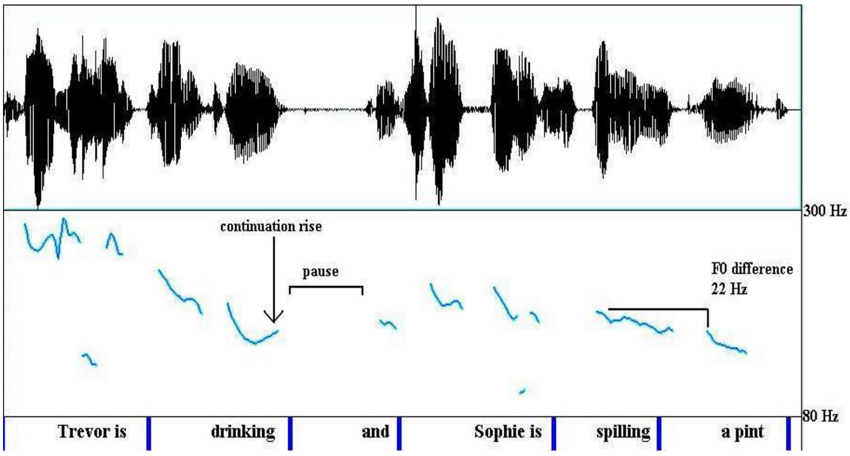

Figure 1: Typical pitch track for non-elliptical condition

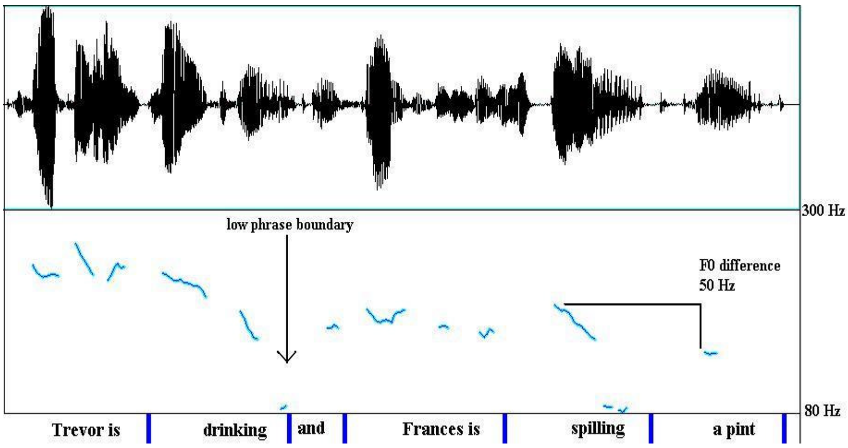

Figure 2: Typical pitch track for RNR condition

3.4 Discussion

The acoustic data for the second conjunct were expected and can be interpreted as reflecting the information structural status of the respective condition. Lengthening and relatively high pitch accentuation of VP2 in RNR sentences correlate with the contrastive focus on this element. The objects in these sentences are deaccentuated. The non-elliptical sentences are unmarked with respect to information structure. The whole VP including the object bears broad focus. Thus, according to the nuclear stress rule, the object is accented.

The data for the first conjunct also reveal a significant prosodic difference between the two conditions. While the non-elliptical sentences exhibit a strong IP boundary with continuation rise, in RNR sentences the boundary between the conjunct is much weaker with virtually no pause. This contrast cannot easily be attributed to information structural differences in RNR vs. non-elliptical sentences, since there is no clear difference in focus distribution in the first conjunct. However, there is an obvious syntactic contrast between the conditions at the end of the first conjunct: while in non-elliptical sentences, VP1 is intransitive, it is a transitive verb in RNR, although the object remains unexpressed. We can, therefore, assume that the prosodic contour of the first conjunct is a reflection of the ellipsis.

An explanation for the interplay between syntactic structure and prosody can be found in phrasing tendencies postulated by Watson & Gibson (2004). According to their model, IP boundaries are determined by the amount of phonetic material to be expressed within an utterance. Speakers have to pause after a certain amount of phonetic material in order to plan the following phrase. This condition is constrained by the syntactic environment. One such constraint concerns syntactic heads and their respective complements. It is strongly favoured that head and complement are included in the same IP. The shorter phrase break in the RNR condition can be explained by assuming that the speaker avoids a strong IP boundary between the head (VP1) and its complement (object NP) in the second conjunct.[3]

4 Speech Perception

The speech perception experiment was conducted to evaluate the influence of prosody on sentence comprehension and especially on the capacity to recover the ellipsis. The first aim was to demonstrate that listeners can use prosodic cues on-line to distinguish the non-elliptical reading from the RNR version.

[3] This account is not unproblematic: Assuming that RNR-type ellipsis is a process of deletion of phonetic material, the object NP in the second conjunct is not analysed to be the complement of VP1.

Secondly, having discovered that the speaker displays specific prosodic renderings not only on the second, but also on the first conjunct, the following question arises: does the prosodic contour on the first conjunct play a role for listeners in parsing the respective reading?

4.1 Method

The experiment is designed to elicit behavioural data that reflect listeners' capacity to use the critical prosodic cues on-line to recover the relevant reading. The experimental paradigm is a cross-modal, binary forced-choice task, in which test persons had to decide whether a written context matches the auditorily presented test sentence.

To allow for an evaluation of the influence that the prosodic features on the first conjunct have on sentence comprehension, it is not sufficient to compare listeners' reaction to the two natural conditions. Both sentence types differ with respect to the first as well as the second conjunct. Therefore, a possible difference in listeners' reaction to the sentence types is not unambiguously attributable to the prosody of the first conjunct. To isolate the influence of the first conjunct, two further conditions were created, so that the prosodic features of both conjuncts were distributed crosswise across conditions.

Cond. 1: Non-elliptical (N=19)	[Nina is RIDing ↑ #] [and Ian is fixing a BIKE]
Cond. 2: RNR (N=19)	[Nina is RIDing ↓] [and Ian is FIXing a bike]
Cond. 3: Non-elliptical splice[4] (N=18)	[Nina is RIDing ↓] [and Ian is fixing a BIKE]
Cond. 4: RNR splice[4] (N=18)	[Nina is RIDing ↑ #] [and Ian is FIXing a bike]
Key: ↑ = high boundary tone, *continuation rise* ↓ = low boundary tone # = final lengthening + pause	

Table 2: Prosodic features of test conditions

Using digital audio software (CoolEdit), the sentences were split into single conjuncts, which were then respliced. Table 2 schematically depicts the crosswise distribution of the conjuncts and their respective prosodic features in the four resulting test conditions. All sentences were normalised to approximately equal loudness.

For each test sentence, a context was devised that either fitted the non-elliptical or the RNR reading but not both. The distribution of context and condition with the expected answer is exemplified in table 3.

[4] The prosodic structure on the second conjunct determines the name of the splice conditions.

Nina is riding and Ian is fixing a bike.	Expected answer in case of	
Context	Non- elliptical sentence (Cond. 1+3)	RNR sentence (Cond. 2+4)
Nina enjoys a ride on her horse. Ian's bike has a flat tyre. He is repairing it in the backyard.	YES	NO
Nina has skipped the riding session and is now cycling to the city centre. Ian's bike has a flat tyre. He is repairing it in the backyard.	NO	YES

Table 3: Context- test sentence pairing with expected answer

In addition to the test sentences, the same amount of filler sentences (spoken by the same speaker) and relevant contexts were provided for the experiment.

Twelve healthy subjects (age range 20-32) with neither hearing nor vision impairments took part in the experiment. All participants were native speakers of English and naive to the aim of the experiment.

4.2 General Procedure

The subjects were seated in a sound-proof room in front of a computer screen including a keyboard and were equipped with headphones. Upon pressing the spacebar, the written context appeared on the screen. Having read and understood the context, the participant was asked to press the spacebar key. The relevant sentence was immediately played via headphones. Directly after auditory presentation, subjects had to judge the appropriateness of context and test sentence as quickly as possible by pressing either the YES or the NO key on the keyboard. A timeout was programmed so that only those answers were recorded that were given within 3500ms after the end of the auditory sentence presentation. Finally, subjects were asked to rate the sentences on a naturalness scale (1 = natural sentence – 5 = unnatural sentence) by pressing the appropriate number key.

The test and filler sentences with respective contexts were presented in randomised order using DMDX presentation software (Forster & Forster 2003). All answers, reaction times and naturalness values were recorded and subjected to statistical analysis. Error rates and reaction times were calculated for each of the four test conditions.

4.3 Results

The error rates demonstrate that the natural RNR sentences were judged correctly below 50% of the time, reflecting chance performance within timeout. In all other conditions, subjects perform significantly above chance. An one-factorial ANOVA reveals a main effect for the factor “sentence type”. Performance on the natural RNR condition differs significantly from all other conditions (multiple comparisons: Bonferroni $p<0.001$). Figure 3 depicts the average percentage of correct answers for each condition.

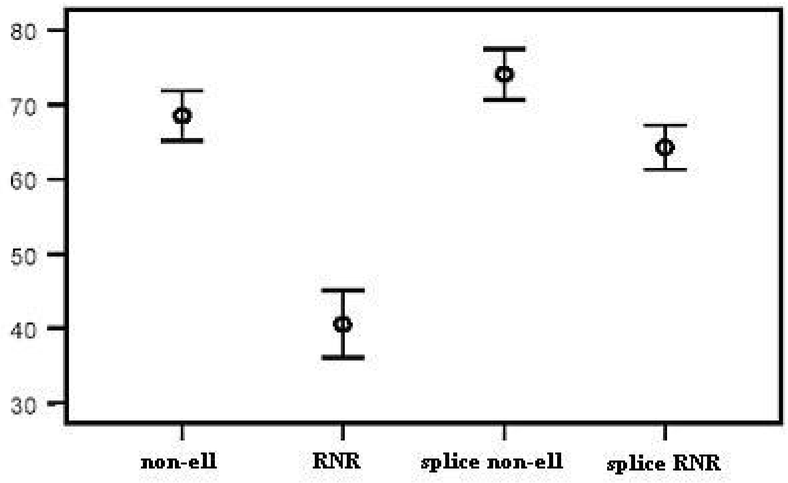

Figure 3: Mean percentage (+/- 1SE) correct responses in test conditions

As to the reaction times, only those values were included into statistical analysis that stem from correct answers. Extremely high reaction times (>3000ms) were also excluded from further analysis. To evaluate the different effects that the factors “sentence type”, “subject” and “answer” (YES vs. NO) had on the reaction times, a 4*12*2 between-subject ANOVA was employed. The results exhibit a pattern comparable to the error rates: reaction times for the natural RNR sentences were distinctly higher than reaction times of the other conditions, which is confirmed by a significant main effect for the factor “sentence type”. Multiple comparisons confirm a significant difference between natural RNR sentences and both splice versions.

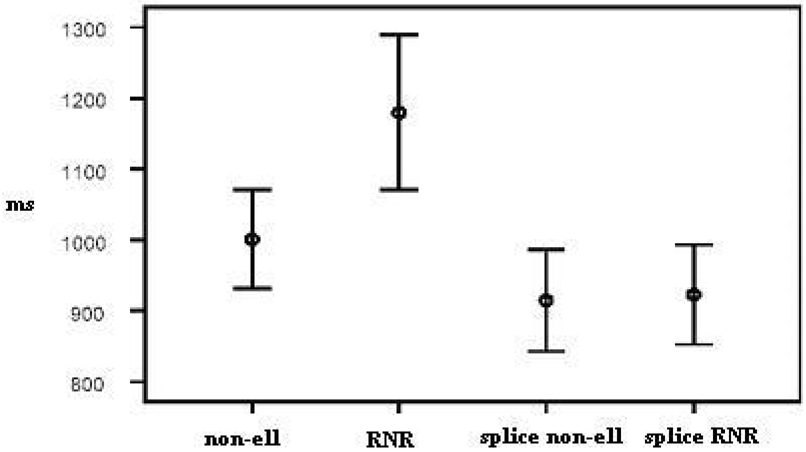

Figure 4: Mean reaction times (+/-1 SE) for correct responses in test conditions

4.4 Discussion

The results of all parameters point to the same direction: the naturally spoken RNR sentences are not easily interpreted and are more difficult to process than the three other conditions.

A comparison of the two natural conditions (non-elliptical vs. RNR sentences) suggests that listeners cannot reliably assign the appropriate reading to RNR sentences within the time frame of this experiment. They perform within chance range. Given this result, the specific prosody of the RNR sentences does not seem to have a directly supporting influence on on-line sentence processing. Rather, this result may be in line with *syntax first* theories of sentence processing that assume processing advantage for simple syntactic structures over more complex structures, irrespective of their prosodic or semantic environment (Frazier 1987). According to this theory elliptical sentences are assumed to be generally more difficult to process than simple non- elliptical sentences with the same word order.

A more detailed analysis of the results gives reason to doubt this interpretation. A direct comparison of listeners' performance on the non-elliptical sentences on the one hand and the spliced RNR sentences on the other does not reveal any difference with respect to error rates and reaction times. The fact that listeners' performance on natural non-elliptical sentences is indistinguishable from spliced RNR sentences is not expected in a *syntax first* perspective. Obviously, listeners were able to assign the appropriate reading to the spliced RNR sentences. The only way to do so was by exploiting the specific prosodic contour that distinguishes RNR from non-elliptical sentences with the same word order. Therefore, under favourable circumstances, prosody does inform the parser early in sentence processing, thereby neutralising difficulties that arise from syntactic complexity. On the other hand, as seems to be the case in the natural RNR condition, sub-optimal prosody[5] hinders sentence processing. In those instances syntactic complexity comes into play as a decisive factor influencing processing ease.

With regard to the question whether the prosodic features on the first conjunct have a role to play in the interpretation of RNR- sentences, the comparison between listeners' performance on natural RNR and spliced RNR-sentences is meaningful. The opposition of these sentence types represents the critical feature, since natural and spliced versions differ only with respect to the prosodic rendering of the first conjunct. The listeners exhibit a significant difference in performance on these sentence types. It is, therefore, safe to conclude that prosodic structure at the end of the first conjunct matters for processing of RNR sentences.

5 General Discussion

The experiments described above yield two main results:

First, the speech production experiment reveals that different prosodic contours are employed for RNR sentences as opposed to non-elliptical sentences with the same word order. The difference concerns the accent position in the second conjunct, reflecting the different focus structure of the two conditions. In addition, the phrase break at the end of the first conjunct (which corresponds to the ellipsis site in RNR sentences) exhibits a marked difference with regard to prosody in the two conditions. The IP boundary is distinctly stronger in non-elliptical sentences than in RNR sentences.

Second, the results of the speech perception experiment substantiate the hypothesis that critical prosodic cues are exploited on-line during sentence processing in order to assign the appropriate reading to an otherwise ambiguous structure. To this end, listeners already use prosodic information at the end of the first conjunct. Listeners' performance demonstrates that processing obstacles due to syntactic complexity can be neutralised in the case of optimal prosodic phrasing. In the case of a strong IP boundary between the two conjuncts in RNR sentences, processing performance is indistinguishable from performance on simple non-elliptical sentences.

Prosodic phrasing in production and perception

It seems reasonable to suppose that the prosodic difference between RNR and non-elliptical sentences correlates with the presence of an ellipsis. In RNR

---

[5] In this case, the comparatively weak phrase break after the first conjunct is sub-optimal for the interpretation of RNR sentences.

sentences the complement of VP1 is not expressed in the same phrase, but its identical counterpart appears at the right periphery of the following phrase. According to syntactic constraints on phrasing, a strong prosodic boundary between VP1 and its complement is strongly disfavoured. This might give rise to the comparatively weak phrase boundary between the conjuncts in RNR as opposed to non-elliptical sentences.

With regard to speech perception of RNR sentences, however, a strong IP boundary is preferred over a weak one by listeners. Thus, RNR sentences obviously represent a structure in which production and perception preferences conflict with regard to the distinctness of the IP boundary.

The role of accentuation in speech perception

As was shown above, the specific information structural environment with contrastive focus on the pre-elliptical element (VP2) in the second conjunct is a constitutive feature of RNR. The narrow contrastive focus is represented by pitch accent on VP2 and deaccentuation of the object in these sentences. However, it seems that the prosodic prominence pattern on the second conjunct is not sufficient in order for listeners to assign the appropriate reading within the time frame of the perception experiment. This is demonstrated by the chance performance that listeners exhibit when confronted with the natural RNR sentences.

This phenomenon can be interpreted with reference to an ERP experiment conducted by Hruska et al. (2001). The authors show that, under certain circumstances, listeners ignore pitch accents. Given that some accents do not provide information necessary for sentence interpretation, such accents are deemed “superfluous” and ignored by the parser. With regard to these findings, the following explanation of listeners’ performance on the different RNR versions is possible. The pitch accent on VP2 in the natural RNR condition seems to remain unrecognised by the parser. The contrastive accent on VP2, however, is a decisive feature of RNR sentences. Ignorance towards this accent leads to the failure to reliably parse the appropriate reading.

Apparently the pitch-accented VP2 is underspecified with respect to its focus type, as long as additional prosodic information on the object is not yet processed. Only relative deaccentuation of the object unambiguously signals that VP2 is contrastively focused. If not forced otherwise, listeners compute unmarked, broad presentational focus on VP2 and thus expect nuclear stress on the object. Only upon encountering the deaccentuated object do listeners have to give up the preferred simple, non-elliptical reading and reanalyse the focus structure in favour of the RNR reading. This reanalysis is reflected in the high processing costs listeners show when confronted with the natural RNR condition.

Spliced RNR versions (which display the same accent pattern on the second conjunct) do not pose such problems for listeners. Correspondingly, it seems as if the specific prosodic rendering of the first conjunct is decisive for the correct interpretation of the accent pattern on the second conjunct. The IP boundary at the end of the first conjunct in the spliced RNR sentences is marked by final lengthening, a high boundary tone and pause. The high boundary tone at the end of the first conjunct together with the pitch fall after VP2 forms a bridge or hat contour, a prosodic feature typical of contrast in Germanic languages (e.g. Wunderlich 1991, van Hoof 2003). English listeners obviously take advantage of this global prosodic parameter to represent contrastive accent on-line on VP2, thereby avoiding costly reanalysis.

The role of prosody in sentence processing

The present experiment shows clearly that prosody informs the parser during on-line sentence processing. Listeners smoothly bypass possible parsing difficulties, which might arise due to syntactic complexity, by recourse to prosody. However, prosody does not always support sentence processing. Sub-optimal prosody might lead the listeners *up the garden path*, forcing them into costly reanalysis. These results are in line with assumptions of Speer et al. (1996) on the role that prosody plays in on-line sentence processing (see also Schafer 1997). In the case when syntactic and prosodic phrase boundaries coincide, processing is trouble free. Problems arise whenever prosodic and syntactic structure conflict, leading to processing effects that reflect syntactic complexity. This idea is supported by the results of the perception experiment. It provides evidence for a strong interaction of prosodic phrasing and accentuation in the course of parsing. Successful processing of pitch accents that correspond to instances of contrastive focus depends on their integration into a global prosodic contour, and especially on the distinctness of phrase boundaries.

It remains to be specified why and under what circumstances listeners' and speakers' preferences regarding the prosodic contour conflict. Furthermore, it is as yet unclear, why listeners seem to ignore the pitch accent on the second conjunct in the natural RNR condition but not in the spliced RNR condition. Which prosodic properties of the first conjunct are responsible for this dissociation?

6 References

Albritton, D.W., G. McKoon and R. Ratcliff (1996): Reliability of prosodic cues for resolving syntactic ambiguities. Journal of Experimental Psychology: Learning, Memory and Cognition 22, 714-735

Alter, K., I. Mleinek, T. Rohe, A. Steube and C. Umbach (2001): Kontrastprosodie in Sprachproduktion und Perzeption. Linguistische Arbeitsberichte 77, 59-79

Beckman, M.E. (1996): The parsing of prosody. Language and Cognitive Processes 11, 17-67

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

Chomsky, N. and M. Halle (1968): The sound pattern of Engish. New York: Harper and Row

Cutler, A., D. Dahan and W. van Donselaar (1997): Prosody in the comprehension of spoken language: A literature review. Language and Speech 40, 141-201

Ferreira, F. (2002): Prosody. In: Encyclopedia of Cognitive Science. London: Macmillan Reference

Féry, C. and V. Samek-Lodovici (2006): Focus projection and prosodic prominence in nested foci. Language 82.1.

Forster, K. and J.C. Forster (2003): DMDX: A windows display program with millisecond accuracy. Behavior Research Methods, Instruments and Computers 35, 116-124

Frazier, L. (1987): Theories of sentence processing. In: J.L. Garfield (ed.): Modularity in knowledge representation and natural language understanding. Cambridge, MA: MIT Press, 291-307

Hartmann, K. (2000): Right node raising and gapping: Interface conditions on prosodic deletion. Amsterdam: John Benjamins

Hruska, C., K. Alter, K. Steinhauer and A. Friederici (2001): Misleading dialogues: Human`s brain reaction to prosodic information. Proceedings of the ORAGE-conference. Aix-en-Provence, 425-430.

Ladd, R. (1996): Intonational phonology. Cambridge: Cambridge University Press

Marcus, M and D. Hindle (1990): Description theory and intonation boundaries. In: G.T.M. Altman (ed.), Cognitive models of speech processing: Psycholinguistic and computational perspectives. Cambridge (MA): MIT Press, 483-512

Nespor, M. and I. Vogel (1986): Prosodic Phonology. Dordrecht: Foris

Pynte, J. and B. Prieur (1996): Prosodic breaks and attachment decisions in sentence parsing. Language and Cognitive Processes 11, 165-191

Schafer, A. (1997): Prosodic parsing. The role of prosody in sentence comprehension. Doctoral dissertation. Amherst (MA): University of Massachusetts

Selkirk, E. (2002): Contrastive FOCUS vs. presentational *focus*: Prosodic evidence from right node raising in English. In: Speech Prosody 2002: Proceedings of the first international conference on speech prosody. Aix-en-Provence, 643-646

Snedeker J. and J. Trueswell (2003): Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. Journal of Memory and Language 48, 103-130

Speer, S.R., M.M. Kjelgaard and K.M. Dobroth (1996): The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. Journal of Psycholinguistic Research 25, 247-268

Terken, J. and S.G. Nooteboom (1987): Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. Language and Cognitive Processes 2, 145-163

Toepel, U. and K. Alter (2004): On the independence of information structural processing from prosody. In: A. Steube (ed.): Information structure. Theoretical and empirical aspects. Berlin: de Gruyter

Van Hoof, H. (2003): The rise in the rise-fall contour: does it evoke a contrastive topic or a contrastive focus? Linguistics 41, 516-543

Warren, P., E. Grabe and F. Nolan (1995): Prosody, phonology and parsing in closure ambiguities. Language and Cognitive Processes 10, 457-486

Watson, D. and E. Gibson (2004): The relationship between intonational phrasing and syntactic structure on language production. Language and Cognitive Processes 19, 713-755

Wilder, C. (1997): Some properties of ellipsis in coordination. In: A. Alexiadou and T.A. Hall (Eds.): Studies on universal grammar and typological variation. Amsterdam: John Benjamins, 59-107

Wunderlich, D. (1991): Intonation and Contrast. Journal of Semantics 8, 239-251