Nuclear Accent Placement and Other Prosodic Parameters as Cues to Pronoun Resolution*

Ekaterina Jasinskaja, Ulrike Kölsch, and Jörg Mayer

University of Potsdam

This paper investigates the influence of prosody on the interpretation of anaphoric pronouns, concentrating especially on the effect of nuclear accent placement. It is well-known that accented and unaccented pronouns generally have different resolution preferences, but it is less obvious that pronoun interpretation can be affected by almost any manipulation of the accentual pattern of the sentence in which the pronoun occurs, even by a manipulation that does not involve the pronoun. However, the latter follows from theories of accentuation such as [1] and in this paper we present experimental support for this prediction. Our results corroborate the view that the influence of accent on pronoun resolution should be derived from a general theory of focus interpretation, rather than from rules defined specifically for accents occurring on pronouns.

We start in Sect. 1 by presenting some background on accentuation and its impact on pronoun resolution. Since accent is a way of signalling contrast, and contrast in turn can be viewed as a rhetorical relation, constraints on pronoun resolution that result from rhetorical structure should be taken into account as well, which is done in Sect. 2. Section 2 also introduces hypotheses related to other prosodic parameters (pitch range and pause duration) which are known to be able to convey aspects of rhetorical structure. Finally, Sect. 3 describes our experiment, and Sect. 4 discusses the results.

1   Accent Placement and Pronoun Resolution

It is well-known that accentuation affects the resolution preferences of anaphoric pronouns. In particular, the effect of accenting the pronoun itself has been studied quite extensively and is illustrated by the following example (coindexing indicates coreference relations):

(1)   a. Paul_i called Jim_k a Republican.
         Then he_i insulted him_k.
      b. Paul_i called Jim_k a Republican.
         Then HE_k insulted HIM_i.

Recent studies have argued that there is nothing special about the role of accent when placed on a pronoun; rather, the effects of accent on pronoun resolution should be derived from a general theory of accentuation and focus semantics [1, 2]. In particular, it is proposed that certain accentual patterns, including those that involve accented pronouns, require the presence of a *contrasting alternative* in the context or a possibility to accommodate such an alternative [3–5]. A contrasting alternative in this case is a constituent (often, a clause) that differs from the clause in question only in the focused ($\approx$ accented) subconstituent(s). Thus in (1b), the accents on the pronouns and the absence of accent on the verb *insulted* are licensed only if the sentence is taken to contrast with the preceding sentence *Paul called Jim a Republican*; this implies that (a) HE must be contrasted with, i.e. distinct from, *Paul* (hence HE $\mapsto$ Jim), (b) HIM must be contrasted with, i.e. distinct from, *Jim* (hence HIM $\mapsto$ Paul), and (c) *insulted*, since it is deaccented, must be viewed as "parallel" to *called a Republican*, so that calling someone a Republican has to be accommodated as a kind of insult.

Most of the existing theoretical analyses of accented/stressed pronouns, notably [6] and [4], seem to treat accent as an independent property of the pronoun.[1] However, theories of accent placement such as [7] or [1] suggest that the decision to accent or deaccent a pronoun is not independent from the decision to accent or deaccent other constituents in the sentence. Thus, for instance, placing no accent on the pronouns in the second sentence of (1a) means almost automatically that the verb *insulted* must be accented. In this paper we present further support for the idea that the dependence of pronoun resolution on accentuation is a by-product of the general functioning of prosodic focus as a contrast-signalling device; however, unlike the previous studies, we would like to emphasise the importance of the overall accentual pattern of a sentence. That is, it does not only matter whether the pronoun is accented or not, but as predicted by Schwarzschild [1], any occurrence of an accent in the sentence, as well as any occurrence of deaccenting is potentially relevant for determining an antecedent. We present below the results of an experiment which show that this prediction is indeed borne out. There is a well-known asymmetry between nuclear and pre-nuclear accents in marking given information, and we will therefore follow Venditti et al. [8] in restricting our attention to the placement of *nuclear* accents—the most prominent, and usually the last, accent in a prosodic phrase.

To illustrate the prediction in question, consider the German example in (2) as well as its English translation in (3):

---

[1] This approach makes it look as if the opposition of stressed and unstressed pronouns behaves like the opposition of strong and weak pronouns, e.g. *it* vs. *that* in English, *er* vs. *der* in German. The latter indeed applies specifically to pronouns, in that the strong/weak pairs often have to be defined in the lexicon, rather than following a productive pattern, whereas accentuation is completely productive in languages like English and German, and is not restricted to pronouns. Although the choice between a strong and a weak pronominal form might not be completely independent from stress, the oppositions are *a priori* distinct and a uniform analysis should be empirically justified.

(2) a. Johann hat die Möhren geschnitten.
   Johann has the carrots cut
 b. Marek hat indes   die Kartoffeln geschält.
   Marek has meanwhile the potatoes   peeled
 c. Außerdem hat er die Kartoffeln geschnitten.
   besides   has he the potatoes   cut

(3) a. Johann cut the carrots.
 b. Meanwhile, Marek peeled the potatoes.
 c. Besides, he cut the potatoes.

The most natural interpretation of the pronoun *er* 'he' in (2c)/(3c) is *Marek*, the only male individual mentioned in the immediately preceding sentence, while the most natural pronunciation of (2c) is with a nuclear accent on the verb *geschnitten* 'cut', indicated by small caps in (4), whith the direct object *die Kartoffeln* 'the potatoes' deaccented.

(4) Außerdem hat er die Kartoffeln GESCHNITTEN
   besides   has he the potatoes   cut

This pattern is explained straightforwardly if we assume that only the previous sentence is relevant for establishing the contrast relation. In that case, the transitive relation *He/Marek X-ed the potatoes* is given, while the verb *geschnitten* 'cut' is contrasted with *geschält* 'peeled,' so the verb is narrowly focused and receives the nuclear accent. However, if both context sentences (2a) and (2b) are taken into account, the question arises, with which of them (2c) should be contrasted. This choice is essential for determining the accentual pattern, and as it turns out, it interacts in a crucial way with the choice of antecedent for the pronoun.

Suppose, as before, that *er* 'he' in (2c) refers to *Marek*, but (2c) is contrasted with (2a). Then the verb *geschnitten* 'cut' is given, but its arguments are contrasted: *er/Marek* with *Johann* and the potatoes with the carrots. So a contrast/givenness-based theory predicts accents on *er* and *Kartoffeln*, cf. (5a). Now suppose that the pronoun refers to the antecedent farther away—*Johann*. If (2c) is contrasted with (2b), the sentences differ in who did what to the potatoes, so accents are expected on the pronoun *er/Johann*, contrasting with *Marek*, and *geschnitten* 'cut,' contrasting with *geschält* 'peeled,' cf. (5b). If, in turn, (2c) is contrasted with (2a), the open proposition *He/Johann cut X* is given and only the objects of cutting, the potatoes and the carrots, are contrasted. Therefore, we predict an accent on *Kartoffeln* 'potatoes,' cf. (5c), whereas the pronoun receives no accent.

(5) a. Außerdem hat ER [Marek] die KARTOFFELN geschnitten
     besides   has he   the potatoes   cut
   b. Außerdem hat ER [Johann] die Kartoffeln GESCHNITTEN
     besides   has he   the potatoes cut
   c. Außerdem hat er [Johann] die KARTOFFELN geschnitten
     besides   has he   the potatoes   cut

This short sketch shows that theories of accent placement based on the notions of contrast and givenness such as [5] and [1] predict a rather complex interplay between pronoun resolution possibilities and accentuation patterns. Since coreference relations play a role in identifying the parallel part of the contrasting clauses, the choice of pronoun antecedent does not only determine whether the pronoun should be accented or not, but also imposes constraints on which other parts of the sentence may be accented. Conversely, one would expect that in discourses like (2) the shift of accent between the direct object and the verb in (2c) should correlate with a change of antecedent for the pronoun *er* 'he.' This study concentrates specifically on the contrast between (4) and (5c) where the pronoun remains unaccented in both versions. Here, the nuclear accent on *geschnitten* 'cut' is expected to correlate with the resolution of *er* 'he' to *Marek*. Resolution to *Johann* would, of course, be dispreferred due to distance considerations, however a nuclear accent on *Kartoffeln* should favour this suboptimal resolution. Testing this hypothesis is the main goal of the experiment presented below, but first a few words on some further corollaries of this hypothesis.

2   On the Role of Discourse Structure

If the above theory is correct, then placement of accent can influence which part of the context a current sentence is contrasted with. Thinking of contrast as a rhetorical relation, along the lines of Mann & Thompson [9] or Asher & Lascarides [10], accent placement can thus affect the *attachment site* of the current sentence in the discourse structure: with the accentual pattern in (4) the sentence is attached with a contrast relation to the immediately preceding sentence; with the accent on the direct object as in (5c), the sentence is attached higher up in the discourse structure, to a sentence that is farther away. In other words, the latter case is an instance of *discourse pop*.

The present work is part of a larger study on prosody as a cue to discourse structure. There is a substantial body of research on discourse prosody (see e.g. [11–18]) showing that pitch range—the fundamental frequency span between the realizations of high and low tones—is higher at the beginning of a discourse unit (e.g. a paragraph) and lower at its end. A switch from one discourse unit to another (topic shift, or discourse pop) is therefore associated with a perceptible reset of pitch range back to higher and larger F0 span. Similarly, discourse pops correlate with relatively longer pauses between utterances [19, 20].

Furthermore, it is well-known that the hierarchical organisation of discourse (global topic structure) affects anaphora resolution. Although, in general, refer-

ents mentioned in more recent sentences tend to be more accessible for pronominal reference, a discourse pop can change this. If a less recent antecedent is related to a more global discourse topic, it can become more salient when that topic is reactivated after the pop. A more recent antecedent, on the other hand, can become less salient, if it is only locally important in the discourse segment just closed off.[2] Consequently, prosodic features signalling a discourse pop are expected to facilitate resolution of pronouns to less recent antecedents, which is supported by our previous experimental studies [29, 30].

Applying these findings to example (2) above, one would expect that a pitch reset in the last sentence, as well as a long pause before it, should favour high attachment to (2a) with corresponding resolution of *er* 'he' to *Johann*. Lack of pitch reset and normal pause length before (2c) should correlate with low attachment to (2b) and resolution of *er* 'he' to *Marek*. In our present experiment the effects of accentuation and global prosodic parameters were studied simultaneously, as we were interested in possible interactions between different prosodic devices signalling discourse attachment.

3 Experiment

3.1 Method

Discourses: For the purposes of the experiment we constructed 40 discourses, each of which consisted of a set of 3 sentences similar to (2), and in which the last sentence could be understood as contrasting with either the first or the second sentence, depending on the interpretation of the pronoun. The potential antecedents were proper names referring to male or female humans (either both male or both female), and always constituted the grammatical subject of the sentence, occurring in sentence-initial preverbal position. Sentence 2 was related to sentence 1 by a discourse adverbial that appeared immediately after the finite verb, cf. *indes* 'meanwhile' in (2b). The target sentences started with a discourse adverbial, cf. *außerdem* 'besides' in (2c), while the ambiguous pronoun *er* 'he' or *sie* 'she,' which was also the grammatical subject, immediately followed the finite verb. We wanted to avoid placing the target pronoun in the absolute sentence-initial position so that the first prenuclear pitch accent could precede it thus enabling the listener to appreciate the pitch range of the utterance before he or she interpreted the pronoun. The nuclear accent in turn always occurred after the target pronoun.

As with (2), all the experimental discourses were designed in such a way that shifting the nuclear accent from one constituent to another in the target sentence would indicate contrast with the first or the second sentence of the context. It should be noted, though, that there is an asymmetry between the accentual

---

[2] This generalisation has been expressed in various forms as the Right Frontier Constraint [21, 10], the stack model [22], the cache model [23], the veins theory [24], the rhetorical distance theory [25], among others, and has been empirically substantiated by e.g. [26–28].

patterns in (4) and (5c). The nuclear accent on the transitive verb as in (4) indicates more or less unambiguously that the verb bears *narrow focus*; that is, the sentence could only be used felicitously as an answer to a question like *What did he do to the potatoes?* or be uttered in a context where the potatoes are given, e.g. if it is contrasted with a sentence that explicitly mentions the potatoes. In contrast, the accent on the direct object in (5c) is ambiguous between narrow focus on *Kartoffeln* 'potatoes' and broad focus on the VP or the whole sentence (cf. e.g. [7, 31]). Thus (5c) can answer both a question like *What did he cut?* and questions like *What did he do?* or *What happened?* Similarly, it can be contrasted with a sentence that only differs from (5c) in the referent of the direct object, e.g. (2a), or with one where, for instance, the whole VP is different: *A: Johann hat die Pfanne gewaschen. B: Nein, er hat [die Kartoffeln geschnitten]*$_{FOC}$. 'A: John washed the frying pan. B: No, he cut the potatoes.' Finally, a transitive sentence with a nuclear accent on the direct object need not be involved in a contrast relation at all and can be uttered "out of the blue," hence this accentual pattern is often called the *default* or the *neutral* pattern.

In order to prevent a confound between the neutral vs. non-neutral accentual pattern distinction and the factor under investigation—contrast with sentence 1 vs. contrast with sentence 2—we made sure that our materials were balanced with respect to whether the "neutral" pattern supported attachment to sentence 1 or 2. To achieve this, half of the discourses were like (2), in that the neutral pattern appeared in the 'contrast with sentence 1' condition, whereas in the other half this was reversed, in that the neutral pattern appeared in the 'contrast with sentence 2' condition. An example of the latter is a discourse like (6) below, cf. the English translation in (7). Here the neutral pattern with the nuclear accent on the direct object *Garten* 'garden' in (6c) appeared in the 'contrast with sentence 2' condition, whereas the marked pattern with the nuclear accent on the verb *gemalt* 'painted' was expected to trigger contrast with sentence 1.

(6) a. Dirk hat den Garten fotografiert.
      Dirk has the garden photographed
     b. Franz hat solange   den Teich gemalt.
      Franz has in that time the pond painted
     c. Dann hat er den Garten gemalt.
      Then has he   the garden painted

(7) a. Dirk took some photos of the garden.
     b. During that Franz painted the pond.
     c. Then he painted the garden.

The syntactic functions of the constituents involved in the accent shift manipulation were varied. There were 12 discourses like (2) and (6) where the nuclear accent was shifted between the (monotransitive) main verb and the direct object. In 8 discourses the accent was shifted between the first and the second object of a ditransitive verb, e.g. *hat Benno ein* Buch *geschenkt* 'gave Benno a book' vs. *hat* Benno *ein Buch geschenkt* 'gave Benno a book,' in (8)/(9).

(8) a. Martha hat Niklas ein Buch überreicht.
       Martha has Niklas a   book   presented

    b. Leonie hat dann Benno eine DVD beschert.
       Leonie has then Benno a   DVD presented

    c. Außerdem   hat sie Benno ein Buch geschenkt.
       apart from that has she Benno a   book   presented

(9) a. Martha gave Niklas a book (as a present).
    b. Then Leonie gave Benno a DVD.
    c. Apart from that she gave Benno a book.

There were 8 discourses in which the accent shift manipulation concerned the direct object of a (mono)transitive verb and a PP- or adverbial modifier of the verb, e.g. HEUTE *ein Seminar versäumt* 'missed a class TODAY' vs. *heute ein* SEMINAR *versäumt* 'missed a CLASS today.' In 2 cases the accent was shifted between a head noun and its PP argument: *ein* BUCH *über Napoleon* 'a BOOK about Napoleon' vs. *ein Buch über* NAPOLEON 'a book about NAPOLEON'; in 4 cases between an NP and its PP modifier: *ein* REGAL *aus Nussbaum* 'a SHELF of walnut wood' vs. *ein Regal aus* NUSSBAUM 'a shelf of WALNUT wood'. Finally, 6 discourses were like (10)/(11) in which the accent was shifted between an NP and its adjectival modifier: *mit einer blonden* AMERIKANERIN 'with a blond AMERICAN' vs. *mit einer* BLONDEN *Amerikanerin* 'with a BLOND American.'

(10) a. Björn tanzte mit einer rothaarigen Amerikaner-in.
        *Björn danced with a.*FEM *red-haired   American-*FEM

     b. Maik tanzte übrigens   mit einer blonden Schwed-in.
        *Maik danced by the way with a.*FEM *blond   Swede-*FEM

     c. Vorher   tanzte er mit einer blonden Amerikaner-in.
        *Before that danced he with a.*FEM *blond   American-*FEM

(11) a. Björn danced with a red-haired American.
     b. By the way, Maik danced with a blond Swede.
     c. Before that, he danced with a blond American.

In the 28 discourses in which the main verb was not involved in the accent shift manipulation, it was important that the verb be part of the background, i.e. that it constitute the parallel (non-contrasting) part of the constrasting sentences. As a result, the verb had to be repeated in all three sentences in a set, e.g. *tanzte* 'danced' in (10), which often led to rather unnatural discourses. To avoid this, in 21 of these 28 discourses, the second and third occurrences of the verb were replaced by synonyms or near-synonyms as in (8) above, where the verbs *überreichen, bescheren* and *schenken* all describe an act of giving a present to someone.[3]

---

[3] Either all the three verbs in a discourse were the same like in (10) or all three were distinct synonyms like in (8). Our intuition was that if the verb of the target sentence were synonymous with the verb of one of the context sentences, but literally repeated

Finally, 42 distractor discourses were constructed. As with the experimental discourses, these consisted of a set of three sentences and mentioned multiple human referents, but varied as to whether or not they contained contrast relations, and as to whether or not the pronouns in sentence 2 or 3 resolved unambiguously on the basis of number and gender features.

Each discourse (experimental or distractor) was accompanied by a *who?*-question of the form in (12) *Who cut the potatoes?* In the experimental items the question was derived from sentence 3 in order to reveal the hearer's interpretation of the pronoun. In distractors, the question addressed any of the three sentences.

(12) Wer hat die Kartoffeln geschnitten?
　　 Who has the potatoes　　cut

Audio Materials: All materials were recorded in an anechoic chamber. The sentences were read by one female speaker in randomised order (i.e. not in the context of the respective discourses), aiming at producing constant pitch range and intensity values. The third sentence of each experimental discourse was recorded in two versions corresponding to the two nuclear accent placements, cf. Figs. 1 and 2.

The sentences were resynthesised and the discourses put together using uniform pitch range and pause duration values following the methodology of Mayer et al. [30]. All signal processing was done using PRAAT [32].

Pitch range was defined as the range between the highest intonationally relevant high tone (HT) and the lowest relevant low tone (LT) within one phrase (sentence). Relevant tones were labelled manually in the original recordings and corresponded usually to high or low tonal targets of pitch accents. For pitch range manipulations, 3 different ranges were defined: normal, compressed and expanded. We determined the normal pitch range of the female speaker as ranging from 150 Hz (baseline) to 270 Hz (topline). Using standard expansion and compression ratios, the expanded pitch range of the speaker was set to 150 Hz baseline and 310 Hz topline and the compressed range to 140 Hz baseline and 250 Hz topline. The first and the second sentence of each discourse were assigned an expanded and a normal range, respectively. Each accentual realization of sentence 3 of the experimental discourses (cf. Figs. 1 and 2) was resynthesised in two versions: once with a compressed pitch range corresponding to the continuity condition and once with an expanded pitch range for the discourse pop condition. Third sentences of distractor discourses were assigned one of the pitch range values (expanded or compressed) on a random basis. Based on the original HT and LT and the target range values, the pitch contour of each sentence was shifted so that the LT was set to the target baseline and multiplied so that the HT reached the target topline.

The original discourses were re-created by concatenating the resynthesised sentences with intervening pauses (intervals of zero amplitude). The standard

---

that of the other, this could have created a bias for attachment to the sentence that contained the literal repetition.
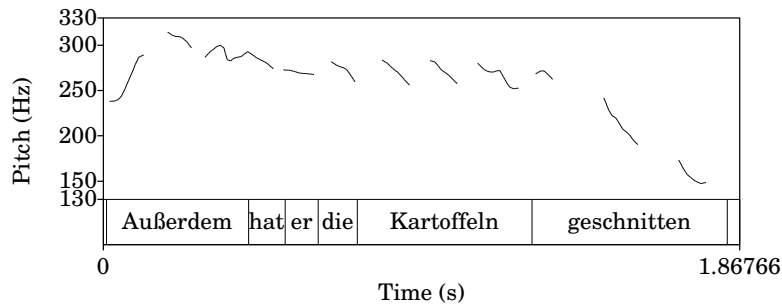
Fig. 1. Pitch track for (4). The falling nuclear accent occurs on *geschnitten* 'cut'.
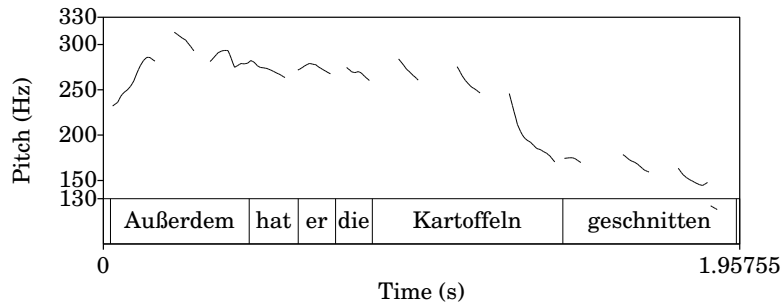

Fig. 2. Pitch track for (5c). The falling nuclear accent occurs on *Kartoffeln* 'potatoes'.

pause length was set to 400 ms. However, in the discourse pop condition an extra long pause of 800 ms was inserted before the last sentence. Figures 3 and 4 show the resynthesised and the reconcatenated realizations of (2) in the continuity and the discourse pop conditions, respectively (the accentual realization of sentence 3 is as in (4), cf. Fig. 1). The horizontal dashed lines indicate the top- and the baselines of the resynthesised sentences. Notice that in the continuity condition (Fig. 3) the pitch range "declines" from the beginning towards the end of the discourse, whereas in the discourse pop condition (Fig. 4) a pitch reset occurs in sentence 3.

In sum, each discourse appeared in four versions corresponding to the four experimental conditions resulting from a 2 by 2 design with accent placement and global prosody (GP) as factors: (1) accent placement in sentence 3 as contrasted with sentence 1 vs. sentence 2; and (2) pitch range of sentence 3 and pause duration before sentence 3 signalling discourse pop vs. discourse continuity.

The final questions were spoken by a male speaker and were appended to the sequences after a silent interval of 1500 ms with the original unmanipulated question intonation.
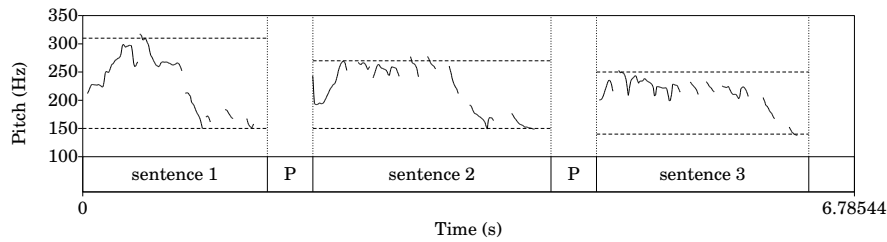
Fig. 3. Prosodic realization of (2) in the continuity condition: the pause between sentence 2 and 3 has standard length (400 ms); sentence 3 has a compressed pitch range.
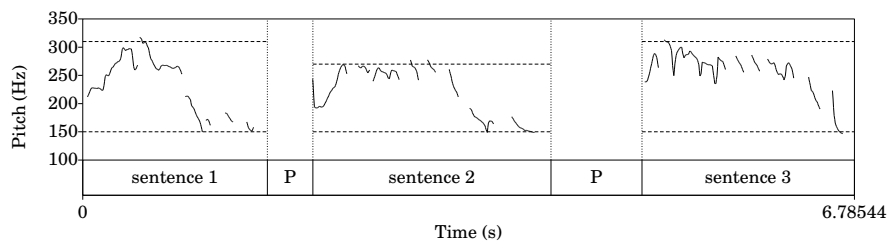

Fig. 4. Prosodic realization of (2) in the discourse pop condition: the pause between sentence 2 and 3 is long (800 ms); sentence 3 has an expanded pitch range.

Procedure: The experimental items were divided into four counterbalanced lists that contained only one version of each item, and mixed with the distractor items. The items were presented in a randomised order. After listening to each item only once, the participants had to answer the questions orally; no choice lists of possible answers were provided. The responses were recorded and classified as indicating resolution to the referent introduced in the first sentence (R1) or the second sentence (R2) or as "incorrect resolution". The response was classified as incorrect if it showed resolution to a referent other than R1 or R2, or if the subject refused to give a definite answer (e.g. by saying *I don't know*).

3.2 Subjects

53 subjects took part in the experiment, all of whom were undergraduate students of linguistics and native speakers of German, and were either paid or received credit for participation. The data of 13 subjects were excluded from the analysis since they failed to give an answer or gave an absurd answer to the test question three or more times. The data of the remaining 40 participants (10 per list) were subjected to statistical analysis.

3.3 Results

The data were aggregated by subjects and by items, the resulting relative frequencies of R1 resolution were square-root arcsine transformed and subjected to ANOVAs. The target pronoun was resolved more frequently (71.8% of times) to the most recent antecedent R2 than to R1 in all conditions, cf. Fig. 5, but there were more resolutions to R1 in the conditions where the accentual pattern corresponded to contrast of sentence 3 with sentence 1 (38.8%) as in (5c), than there were R1 resolutions in the conditions where the accentual pattern corresponded to contrast of sentence 3 with sentence 2 (17.4%) as in (4). The main effect of accentuation was significant both by subjects and by items $[F_1(1, 39) = 32.65, p < .001,$ and $F_2(1, 39) = 59.75, p < .001]$. The main effect of global prosodic parameters was less strong (30.8% vs. 25.4%) and only significant by items $[F_1(1, 39) = 1.66, p = .21,$ and $F_2(1, 39) = 7.15, p < .05]$. We found no interaction between the accentuation and global prosody factors $[F(1, 39) < 1]$.
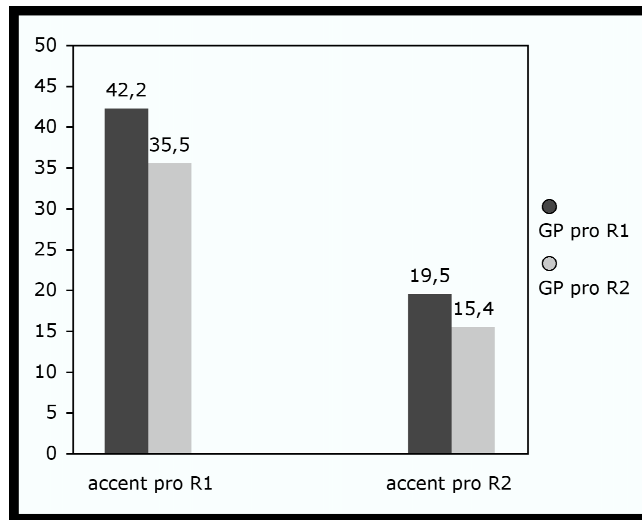

Fig. 5. Number of resolutions of the target pronoun to R1, the antecedent introduced in sentence 1, in %.

4 Discussion and Conclusions

These results corroborate our hypothesis that the placement of nuclear accent can affect pronoun interpretation by determining with which sentence in the context the current sentence should be contrasted. Although, in general, pronoun resolution to the most recent antecedent is preferred, this preference is overridden more often if the accentual pattern of the sentence containing the

pronoun indicates that it should be contrasted with an alternative realized earlier in the discourse, in which case the pronoun (if it is unaccented) is resolved to an antecedent occurring in that alternative. This supports the predictions of theories that claim effects of accent on pronoun resolution to be a by-product of the general theory of accentuation as a contrast-signalling device. However, our results complement those of previous empirical studies in showing that the overall accentual pattern of the sentence also plays a role in determining the contrasting alternative, so that even the accentuation of constituents other than the pronoun can affect its resolution.

Since the effect of global prosody was only significant by items, this effect is more difficult to interpret, as are also the results regarding interaction between global prosody and accentuation as factors. It seems that accentuation and other prosodic parameters may work as independent factors. For this result to be conclusive, however, a stronger main effect of global prosody would need to have been measured. Our previous findings [30] show that the effect of pitch range and pause duration on pronoun resolution is rather subtle (affecting upto 10% of resolutions) and tends to disappear when the discourse pop is not signalled strongly enough, e.g. if different prosodic features do not "cooperate" in indicating a strong prosodic break. This could be one reason why the effect of global prosody in the present experiment was rather weak.Using more strongly expressed prosodic contrasts between the discourse pop and the continuity conditions could help increase the related effect. Another possible explanation for the weakness of the effect is that contrast is a coordinating, or multinuclear, discourse relation [9, 10], and as such is generally thought to assign equal discourse-structural prominence to the sentences it connects. Under this view it is not clear whether the discourse segment that is closed off by the discourse pop in our experimental items (sentence 2) has a subordinated status with respect to sentence 1 or not (see discussion in Sect. 2). However, it is interesting that the global prosody effect that we found is nevertheless in the direction predicted by the Right Frontier Constraint and similar theories: if a pitch reset in the target sentence and a longer pause before it indicate a discourse pop, the pronoun is resolved to an earlier antecedent more frequently than in the continuity condition. This suggests that listeners can sometimes accommodate one of the segments connected by a contrast relation as being discourse-structurally subordinate to the other.

In conclusion, this work contributes to the study of prosody and its interpretation in discourse by demonstrating that pronoun resolution is only one of a whole range of semantic effects of discourse structure conveyed by prosody.

References

1. Schwarzschild, R.: GIVENness, AVOIDF and other constraints on the placement of accent. Natural Language Semantics (1999) 141–177
2. Rooth, M.E.: A theory of focus interpretation. Natural Language Semantics **1** (1992) 75–116

3. Wolters, M., Beaver, D.: What does *he* mean? In Moore, J., Stenning, K., eds.: Proceedings of the Twenty-Third Annual Meeting of the Cognitive Science Society, New Jersey, Lawrence Erlbaum Associates (2001) 1176–1180
4. de Hoop, H.: On the interpretation of stressed pronouns. In Blutner, R., Zeevat, H., eds.: Optimality Theory and Pragmatics. Palgrave Macmillan (2003) 25–41
5. Kehler, A.: Coherence-driven constraints on the placement of accent. In: Proceedings of the 15th Conference on Semantics and Linguistic Theory (SALT-15), Los Angeles, CA (2005)
6. Kameyama, M.: Stressed and unstressed pronouns: Complementary preferences. In Bosch, P., van der Sandt, R., eds.: Focus: Linguistic, Cognitive and Computational Perspectives. Cambridge University Press (1999) 306–321
7. Selkirk, E.: Sentence prosody: Intonation, stress and phrasing. In Goldsmith, J., ed.: Handbook of Phonological Theory. Blackwell, Oxford, UK (1995) 550–569
8. Venditti, J.J., Stone, M., Nanda, P., Tepper, P.: Discourse constraints on the interpretation of nuclear-accented pronouns. In: Proceedings of the 2002 International Conference on Speech Prosody, Aix-en-Provence, France (2002)
9. Mann, W.C., Thompson, S.: Rhetorical Structure Theory: Toward a functional theory of text organization. Text **8**(3) (1988) 243–281
10. Asher, N., Lascarides, A.: Logics of Conversation. Studies in Natural Language Processing. Cambridge University Press (2003)
11. Lehiste, I.: The phonetic structure of paragraphs. In Cohen, A., Nooteboom, S.G., eds.: Structure and Process in Speech Perception. Springer (1975) 195–203
12. Silverman, K.E.A.: The Structure and Processing of Fundamental Frequency Contours. PhD thesis, University of Cambridge (1987)
13. Swerts, M., Geluykens, R.: The prosody of information units in spontaneous monologue. Phonetica **50** (1993) 189–196
14. Sluijter, A.M.C., Terken, J.M.B.: Beyond sentence prosody: Paragraph intonation in Dutch. Phonetica **50** (1993) 180–188
15. Nakajima, S., Allen, J.F.: A study on prosody and discourse structure in cooperative dialogues. Phonetica **50** (1993) 197–210
16. Ayers, G.M.: Discourse functions of pitch range in spontaneous and read speech. Ohio State University Working Papers in Linguistics **44** (1994) 1–49
17. Möhler, G., Mayer, J.: A discourse model for pitch-range control. In: Proceedings of the 4th ISCA Tutorial and Research Workshop on Speech Synthesis, Perthshire, Scotland (2001)
18. den Ouden, H.: Prosodic Realizations of Text Structure. PhD thesis, University of Tilburg (2004)
19. Grosz, B., Hirschberg, J.: Some intonational characteristics of discourse structure. In: Proceedings of the 2nd International Conference on Spoken Language Processing, Banff, Canada (1992) 429–432
20. Swerts, M.: Prosodic features at discourse boundaries of different strength. Journal of the Acoustical Society of America **101** (1997) 514–521
21. Polanyi, L.: A formal model of the structure of discourse. Journal of Pragmatics **12** (1988) 601–638
22. Grosz, B.J., Sidner, C.L.: Attention, intentions and the structure of discourse. Computational Linguistics **12**(3) (1986) 175–204
23. Walker, M.A.: Limited attention and discourse structure. Computational Linguistics **22**(2) (1996) 255–264
24. Cristea, D., Ide, N., Romary, L.: Veins theory: A model of global discourse cohesion and coherence. In: Proceedings of COLING-ACL 1998. (1998) 281–285

25. Kibrik, A.A.: Reference and working memory: Cognitive inferences from discourse observations. In van Hoek, K., Kibrik, A.A., Noordman, L.G.M., eds.: Discourse Studies in Cognitive Linguistics. Benjamins, Amsterdam (1999) 29–52
26. Anderson, A., Garrod, S.C., Sanford, A.J.: The accessibility of pronominal antecedents as a function of episode shifts in narrative text. Quarterly Journal of Experimental Psychology **35a** (1983) 427–440
27. Hitzeman, J., Poesio, M.: Long distance pronominalisation and global focus. In: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Montreal, Canada (1998) 550–556
28. Cristea, D., Ide, N., Marcu, D., Tablan, V.: Discourse structure and co-reference: An empirical study. In: Proceedings of the ACL Workshop on The Relationship Between Discourse/Dialogue Structure and Reference, USA, University of Maryland (1999)
29. Jasinskaja, E., Kölsch, U., Mayer, J.: Global prosodic parameters and anaphora resolution. In Auran, C., Bertrand, R., Chanet, C., Colas, A., Cristo, A.D., Portes, C., Reynier, A., Vion, M., eds.: Proceedings of the International Symposium on Discourse and Prosody as a Complex Interface, Aix-en-Provence, France (2005)
30. Mayer, J., Jasinskaja, E., Kölsch, U.: Pitch range and pause duration as markers of discourse hierarchy: Perception experiments. In: Proceedings of Interspeech 2006, Pittsburgh PA, USA (2006)
31. Wagner, M.: Asymmetries in prosodic domain formation. MIT Working Papers in Linguistics **49** (2005) 329–367
32. Boersma, P., Weenink, D.: Praat: doing phonetics by computer. University of Amsterdam (2005)