The prosody of grouping names in German and Hindi[1]

Caroline Féry and Gerrit Kentner
University of Potsdam

1 Introduction

Prosody reflects syntax, but experimental work is still needed to establish just how closely and how exactly. This paper addresses the prosodic patterns of coordinated nouns grouped in different ways, as produced under experimental conditions. A sequence of nouns like *Anna and Bill or Mary* is an ambiguous structure, in the same way as an arithmetic procedure like *3x2–1*, which can be resolved as 5 or as 3, depending on the order of the operations. Researchers have examined how different groupings of names or numbers are realized prosodically (for instance Lehiste 1983, Raczaszek et al. 1999 and Wagner 2005 for English, and Hunyadi 2006 for Hungarian). All authors focus on differences in duration and find a strong dependency between the duration of constituents and their place in the groupings. In the production experiments reported here, simple groupings with different structures as well as embedded groupings were studied. We measured not only duration, but also the scaling of pitch realization. And two languages with different prosodic grammars are compared, namely German and Hindi. We expect that the general prosodic properties of the two languages are reflected in the way they implement prosody in the rendition of simple and recursive (embedded) groupings.

[1] This paper is part of the project Prosody in Parsing, financed by the DFG. Many thanks are due to Umesh Patil, who prepared the Hindi material and who ran the production experiment in Delhi, as well as to Caroline Magister and Verena Thießen, who ran the experiment for German. We are also grateful to Daniel Quernheim for technical assistance.

2 Background and previous experimental work

There has been a keen interest in the psycholinguistic and phonetic literature as to how prosodic boundaries correlate with syntactic structure, especially in the case of ambiguous structures. Cooper & Paccia-Cooper (1980), Gee & Grosjean (1983), and Ferreira (1993) examine the placement as well as the strength of prosodic and intonational breaks in relationship to syntactic structure in speech production. Carlson, Clifton & Frazier (2001) and Clifton, Carlson & Frazier (2002) discuss the interpretation of prosodic boundaries with respect to sentence processing. See Watson & Gibson (2004) for a summary of previous research.
Watson & Gibson (2004) provide a model of prosodic boundaries called the Left-hand side/Right-hand side Boundary hypothesis (LRB), summarized in (1), in which the size of the preceding and of the following syntactic constituents is the predictor for the strength of prosodic (intonational) boundaries. The material they used consisted of attachment of relative clauses, and is thus very different from that used here.

(1) The Left-/Right-hand side Boundary Hypothesis (LRB, Watson & Gibson 2004)
The likelihood of an intonational boundary at a word boundary is a function of:
(a) the size of the most recently completed constituent and
(b) the size of the upcoming constituent if it is not an argument of the most recent head.

Watson and Gibson's own experiments suggest that the LRB is too simplistic: while the LRB assumes equal influence of constituents to the left and to the right, their results show that the size of the preceding constituent has a much stronger influence on the likelihood of a boundary than the size of the upcoming one (cf. Kentner 2007, who confirms this asymmetry for German).

Wagner (2005) observes that the LRB only predicts effects of adjacent constituents and cannot account for non-local effects of syntactic structure on boundary strength. He proposes an alternative model which relates the strength of prosodic boundaries to syntactic levels of embedding rather than to the size of adjacent constituents. This is the Scopally Determined Boundary Rank (SBR) in (2).

(2) Scopally Determined Boundary Rank (SBR, Wagner 2005):
If Boundary Rank at a given level of embedding is n, the rank of the boundaries between constituents of the next higher level is n+1.

The SBR predicts that the prosodic boundaries between the constituents in (3a) all have the same rank, as all the constituents are at the same level of embedding. But in (3b), the boundary between q and r is stronger than the one between r and s, because r is part of the complex constituent (r and s), i.e. r is more deeply embedded. The first constituent p is at the same level of embedding as q, and both are at the same level as the complex constituent (r and s) in (3b). Thus, the boundary between p and q is predicted to have the same rank as the boundary between q and (r and s). That is, the complexity of (r and s) in (3b) has non-local effects according to the SBR. This property distinguishes the SBR from the LRB, which predicts only local calculation of boundary strength.

(3) a. p or q or r or s

b. p or q or (r and s)

Although the SBR algorithm correctly predicts non-local effects, it cannot easily account for the finding that the boundary strength increases with the size of the upcoming constituent, as predicted by the LRB and as confirmed by the results of both Watson & Gibson and Wagner. Moreover, as Wagner (2005) acknowledges, the SBR's success crucially depends on the use of different normalizing procedures according to the conditions.

Given the problems of the LRB and SBR algorithms, we propose a new account of boundary strength based on two general principles that we call Proximity and Similarity. These principles partly coincide with Watson & Gibson's LRB and partly with Wagner's SBR. However, they avoid the shortcomings of the LRB and the SBR and, at the same time, predict an effect of directionality that is not captured by either of the above accounts. This way, we can account for the fact that right-branching structures appear to be prosodically less articulated than left-branching ones. The two principles and their predictions will be explained in section 3.3.

Our experiment uses different groupings of names and elicits them in a way similar to the one used in Wagner's study. Some differences between this study and ours are that, first, pitch scaling relationships like downstep and upstep are studied in addition to duration. Second, a language with a prosodic system different from that of German (and English) is compared. German is an intonation language, in which pitch accents and boundary tones are varied freely to express pragmatic meanings. We expect that the rich intonational system of this language is also used to indicate the syntactic structure of the groupings. Hindi has a different intonational system. It is a 'phrase language' (Féry 2008, Patil et al. 2008), in which the melody of sentences is primarily determined by invariant phrasal tones. It is thus much more rigid in its use of intonation than German, and we would like to investigate whether this rigidity hinders the prosody from expressing the various syntactic groupings. We thus expect to find differences in the way prosody reflects syntactic grouping, or that different correlates of prosody are used than in German.

In sections 3 and 4, the experiments in German and in Hindi are reported in turn, and section 5 presents a discussion of the results.

3 German production experiment

3.1 Method and material

The material consisted of different groupings of three or four names, all disyllabic and trochaic, like *Mila, Nino* and *Willi*. All groupings tested in the experiment are illustrated in (4), where N1 stands for the first name, N2 for the second name and so on. The conjunction *und* 'and' was always used inside brackets, and the conjunction *oder* 'or' outside of brackets. As can be seen, structures 4.4 and 4.5 include embedded groupings, which are right-branching in the case of 4.4 and left-branching in the case of 4.5.

(4) 3.1 N1 or N2 or N3

3.2 (N1 and N2) or N3

3.3 N1 or (N2 and N3)

4.1 N1 or N2 or N3 or N4

4.2 N1 or N2 or (N3 and N4)

4.3 (N1 and N2) or N3 or N4

4.4 N1 or (N2 or (N3 and N4))

4.5 ((N1 and N2) or N3) or N4

4.6 (N1 and N2) or (N3 and N4)

These structures were inserted in 4 different contexts, one of which is illustrated in (5), with two variants.[2] (5a) triggers the structure 3.2, whereas (5b) elicits the structure 4.3. The other variants were rendered explicit with additional descriptions (see Appendix).

(5) a. Context: Susi and Lena always do sports together, Willi also does a lot of sports. Question: With whom do you want to go for a swim tomorrow?

b. Context: Susi and Lena always do sports together, Rudi is also enthusiastic about sports. And Nino regularly goes to the swimming pool anyway. Question: With whom do you want to go for a swim tomorrow?

The informants heard the question over headphones and saw it on the screen as well. The contexts and the answers in full text were only presented visually. What they had to read was written on the same slide as the context, but consisted only of the names, grouped with the help of parentheses, as exemplified in (6). A logical form was also provided, as shown in (6c).

(6) a. (Susi und Lena) oder Willi.

b. (Susi und Lena) oder Willi oder Nino.

c. $(a \wedge b) \vee c \vee d$

---

2 This procedure was inspired by the one used by Wagner (2005), but the groupings tested were different. Wagner tested only four names in 11 different patterns. We tested three names, as well, in order to have a simpler picture of the effect of groupings in non-embedded patterns.

The speakers were 21 female students at the University of Potsdam, monolingual speakers of German in their twenties, coming from the northern area of Germany. They were paid 6 Euros or got credit points for their participation.

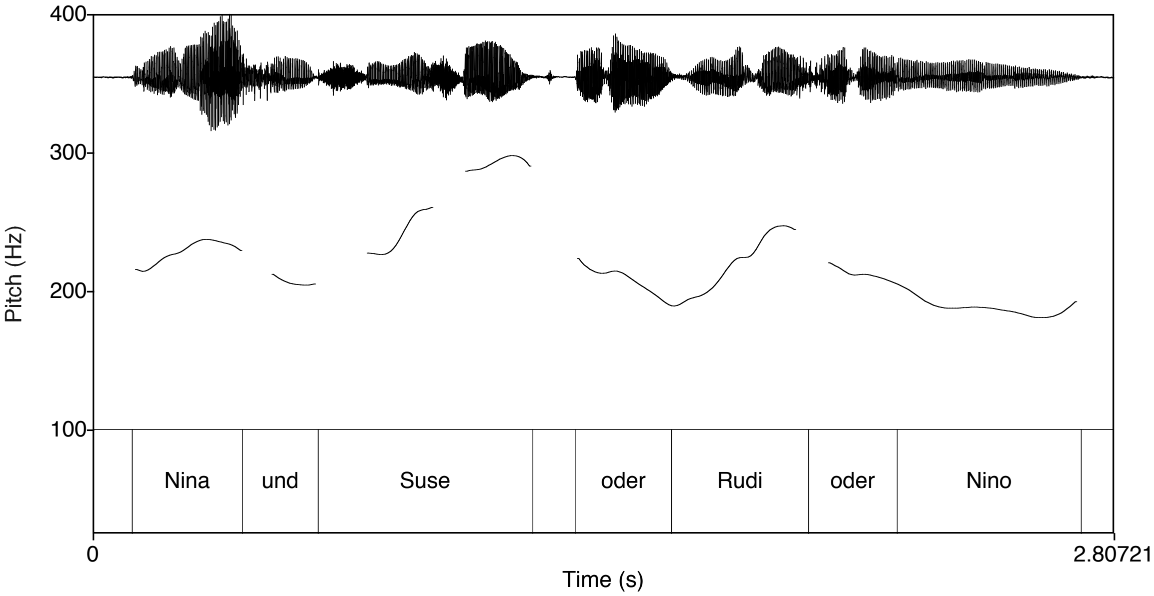
Recordings were made in a soundproof booth on a DAT recorder. The participants were instructed to read the context carefully and to pay attention to the best way of realizing the groupings. They were given as much time as they wanted to utter the answer, and had the opportunity to correct themselves.

Altogether, 756 sentences were recorded and analyzed, 252 with three names (21 subjects x 3 conditions x 4 contexts), and 504 sentences with four names (21 subjects x 6 conditions x 4 contexts).

3.2 Measurements

An example of a realization is given in Figure 1.

Figure 1: pitch track of an answer of the type 4.3



The recordings were re-digitized from DAT at a sampling frequency of 44.1 kHz and 16 bit resolution. Every name as well as every conjunction were labeled and delimited by a boundary set manually in an annotation tier. A break was surrounded by two boundaries. This procedure allowed us to perform semi-automatic measurements with Praat scripts (Boersma & Weenink 2008). Duration was calculated on every item, as well as on the pauses between names and on the conjunctions. We chose to include pauses because we consider them part of the boundaries (see also Gee & Grosjean 1983, Selkirk 1984, Wagner 2005 and Wightman et al. 1992). Thus duration was measured as including the name plus the following pause. However, in a comparison with measurements without pauses, no relevant difference could be observed, except that in Hindi, the differences between the conditions were clearer with the pauses than without them.

The pitch analysis was conducted using a Hamming window of 0.4 seconds length with a default 10 ms analysis frame. The pitch contour was smoothed using the Praat smoothing algorithm (frequency band 10 Hz) to diminish microprosodic perturbations. Time-normalized contours were created by dividing up each constituent into five equal-sized intervals and aggregating the mean F0 (in Hz) over speakers (n=20) and sentences (n=10) for each interval.

All measurements were checked post hoc, and corrected manually when necessary.

Statistic analyses were done using the statistical computing environment R.

3.3 Hypotheses

Based on earlier results from prosody research (Grabe 1998, Truckenbrodt 2002, Féry & Kügler 2008, and others), some assumptions about the production of the expressions can be formulated.

The realizations without grouping, 3.1 and 4.1, are taken as baselines and all other patterns are compared in relation to these baselines. In the baseline patterns without groupings, all names are of equal prominence and are separated by boundaries of the same strength. Every name forms a prosodic phrase (p-phrase) by itself, and no phrase of a higher level is formed on sequences of names. This is illustrated in (7). The entire expression forms an intonation phrase (i-phrase). In such a case, each name gets a pitch accent, which is expected to be rising (L*H) in non-final names and falling (H*L) in final names. L* and H* are the pitch accents, and the trailing tones H and L are the boundary tones of their respective domain. Pitch and duration of the low part of the final constituent are expected to be identical in all cases. In other words, we expect neutralization of the prosodic boundary at the end of all patterns, due to a final low boundary tone at the end of a declarative sentence. Another prediction is that in the default case, every high tone is downstepped relative to the preceding one, and no difference in duration occurs among the names.

(7) L*H L*H L*H H*L
$((\text{N1})_P$ $(\text{or N2})_P$ $(\text{or N3})_P$ $(\text{or N4})_P)_I$

If groupings are reflected in prosody, this happens by means of changed pitch accents, boundary tones and duration, the main intonational events. We derive our hypotheses about the prosodic realization of different syntactic groupings from two general principles that are in line with accounts of the mapping of syntactic and prosodic phrases, as proposed for instance in Selkirk (1984, 1995), Wagner (2005) or Féry & Samek-Lodovici (2006), as well as with more general principles of grouping, as in Hunyadi (2006).

Our expectations are partly in agreement with Watson & Gibson's LRB, which claims that syntactic complexity plays a role in the likelihood of prosodic boundaries, and partly with Wagner's SBR, which posits that embedding of syntactic constituency should be reflected in the strength of the prosodic boundary. However, we add an effect of directionality that is not captured by either of these accounts.

We propose two general principles which operate on the interface between syntactic structure (grouping) and prosodic structure.

First, *Proximity* operates on syntactic constituency, reflecting syntactic boundaries in prosodic and temporal structure. It claims that adjacent elements grouped together into one constituent should be realized in close proximity. A corollary of Proximity is the opposite effect: adjacent elements not grouped together into one constituent should be realized with prosodic distance. This effect is formulated in (8b) as Anti-Proximity.

(8) Proximity

a. Adjacent elements grouped together into one constituent are realized in close prosodic and temporal proximity.

b. (Anti-Proximity): Adjacent elements in separate constituents are realized with prosodic and temporal distance.

Proximity between two elements can be achieved by shortening the duration of the left-hand element and decreasing its pitch accent and boundary tone, thereby *weakening* a group internal boundary. In Anti-Proximity, longer duration of the second element and a higher boundary tone increase the distance to adjacent material to the right that is not part of the same constituent. The boundary is thus *strengthened*.

Note that this implies directionality because it is always the realization of the first of two elements that reflects whether the element to its right belongs to the same constituent or not. In other words, the prosodic realization of a given element only affects its relation to constituents to the right but not to the left. This effect of directionality is also in line with the (uncontroversial) assumption that structure building in speech production is organized incrementally in a left-to-right fashion as proposed by Pierrehumbert (1980), Levelt (1989) and Phillips (2003).

Second, *Similarity* operates on the depth of syntactic and prosodic embedding. It claims that constituents at the same level of embedding should be realized in a similar way, that is, they should be similar in pitch and duration, irrespective of their inherent complexity.

(9) Similarity
Constituents at the same level of embedding are realized in a similar way prosodically, irrespective of their inherent complexity.

The two principles are predicted to interact to shape the prosodic structure of phrases. In case of conflict, they may cancel each other out, but if they are in accord, they amplify each other.

Structures 3.1 and 4.1, without any grouping, are taken as baselines, and serve as points of comparison for the other patterns. According to Similarity, all names should be separated by prosodic boundaries of the same size. In other words, no difference in pitch or duration is expected between these elements.

Grouped elements form p-phrases together, as shown in (10) for 4.4 with an embedded grouping, and these higher levels of prosodic grouping, corresponding to the syntactic structure, should have an effect on tones and duration. We simply assume that prosodic phrasing at the level of the p-phrase is recursive, and do not try to fulfill the non-recursivity assumption of the Strict Layer Hypothesis (Selkirk 1984, Nespor & Vogel 1986).

(10) L*H L*H L*H H*L

$((N1)_P\ ((or\ N2)_P\ ((or\ N3)_P\ (or\ N4)_P)_P)_P)_I$

The predicted effects of Proximity and Similarity are illustrated in (11). All simplex elements at a given level of embedding are subject to Similarity (S). The weak (left) element of a grouping is governed by Proximity (P), and all elements in front of a left or right parenthesis are subject to Anti-Proximity (AP). Proximity has a weakening effect on prosodic boundaries, and both Similarity and Anti-Proximity have a strengthening effect. We do not consider N3 in the three-name groups or N4 in the four-name groups, because of neutralization due to IP finality, which cancels all other effects.

(11) 3.1 N1 or N2 or N3
S S

3.2 (N1 and N2) or N3
P AP

3.3 N1 or (N2 and N3)
S,AP P

4.1 N1 or N2 or N3 or N4
S S S

4.2 N1 or N2 or (N3 and N4)
S S,AP P

4.3 (N1 and N2) or N3 or N4
P AP S

4.4 N1 or (N2 or (N3 and N4))
S,AP P,S,AP P

4.5 ((N1 and N2) or N3) or N4
P AP S,AP

4.6 (N1 and N2) or (N3 and N4)
P AP P

Within groupings, the left-hand elements are weakened, which means that they are expected to be realized with lower pitch and to be shortened (to achieve proximity). The right-hand element of a grouping should lengthen and bear higher high tones, thereby increasing the distance to the material external to the grouping (Anti-Proximity). Likewise, simplex non-grouped elements that immediately precede a grouping should lengthen to increase the distance between the constituents.

Sentences that contain both simplex and complex constituents at a given level of embedding are predicted to display higher tones as well as lengthening of the simplex constituents to achieve similar tonal and durational patterns across that level. At the same time, embedded elements are weakened. As an example, the structures in 4.2 and 4.3 each display one simple grouping of two elements into one constituent. As a result, there are three constituents at the top level in these conditions, two simplex ones and a complex one. The simplex elements are predicted to be adjusted to the complex constituent, both in pitch and in duration, to achieve similarity across constituents at the top level. In addition, the element outside of, but left-adjacent to a grouping is predicted to lengthen to reflect its increased distance to the grouped constituent (Anti-Proximity) (note that this effect only applies to 4.2 as there is no left-adjacent simplex element in 4.3).

To sum up, Proximity has local effects: weakening of the left-hand and strengthening of the right-hand element within a grouping, as well as strengthening of the element left-adjacent to a grouping. Similarity implies that syntactic grouping has non-local effects as well: compared to the baseline, all simplex elements at the top level are strengthened (even ones that are not adjacent to groupings).

3.4 Results for three names

We first compare the realizations 3.2 and 3.3 with the baseline realization 3.1 without brackets. The statistical results are listed in the Appendix if not mentioned explicitly.

First, examine the results for duration in Figure 2. The three names of the baseline pattern 3.1 (black columns) display small differences in duration; the slightly longer length of N2 (mean difference compared to N1 is about 40 ms) is significantly different from N1 ($t=3.8$, $p<0.001$). We return to this effect in the discussion section. Compared to the baseline, pattern 3.2 (dark grey) has a significantly shorter N1 (a group-initial element) and a significantly longer N2 (a group-final element). In contrast, in pattern 3.3 (light grey), N1 (simplex element, left-adjacent to a grouping) is longer while N2 (group-initial element) is shorter than the baseline. We also see that N3's duration is neutralized.

Indeed, this neutralization of the last name is persistent in all conditions, as we will see, both in duration and in pitch.

Figure 2: mean duration in ms of the three names per condition (German).

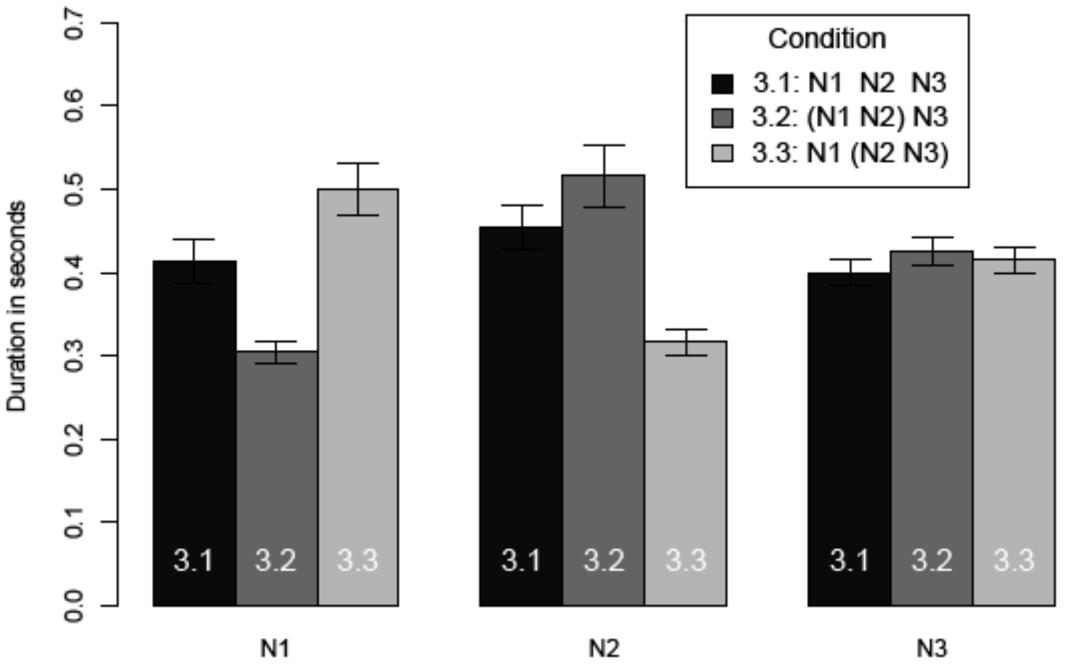
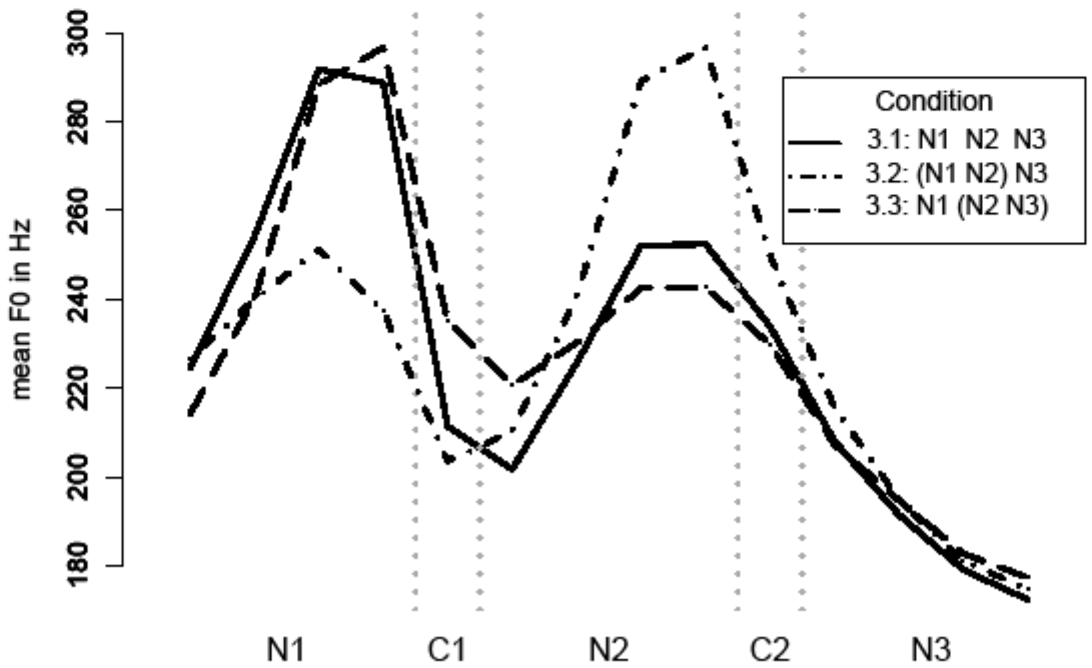


Figure 3 shows the results for pitch. The intervals called C1 and C2 stand for the conjunctions between the names. In the description of the pitch contours, we concentrate on the high tones on the names themselves, and largely ignore the conjunctions, which behave as transitions between the names. The low tones are also discarded in the discussion. These remarks hold for the statistics in the Appendix as well. The baseline pattern 3.1 presents downstep between N1, N2, and N3. The tonal pattern of 3.3, a right-branching structure, is similar to the baseline 3.1. They both have a high N1 and subsequent downstep on the further two names. The N2 of pattern 3.3 is slightly but significantly lowered as compared to the baseline condition 3.1. Pattern 3.2 shows an important difference: the high tone on N1 is much lower than the baseline, while N2 reaches a higher pitch value (upstep). N3 is neutralized in all patterns. Contrary to what was observed for duration, there is no effect for pitch for N1 in 3.3. It is not significantly higher than N1 in the baseline condition 3.1. However, due to the lower pitch on N2, the downstep from N1 to N2 is significantly steeper in this condition. As a result the difference in pitch between N1 and N2 is larger in 3.3 than in 3.1.

Comparing the high tones across conditions, it is conspicuous that the upstepped value of N2 in condition 3.2 has the same height as N1 in the other conditions. And similarly, N1 in 3.3 has roughly the same height as downstepped N2 in the other conditions.

Figure 3: mean pitch in Hz of the three names per condition (German).



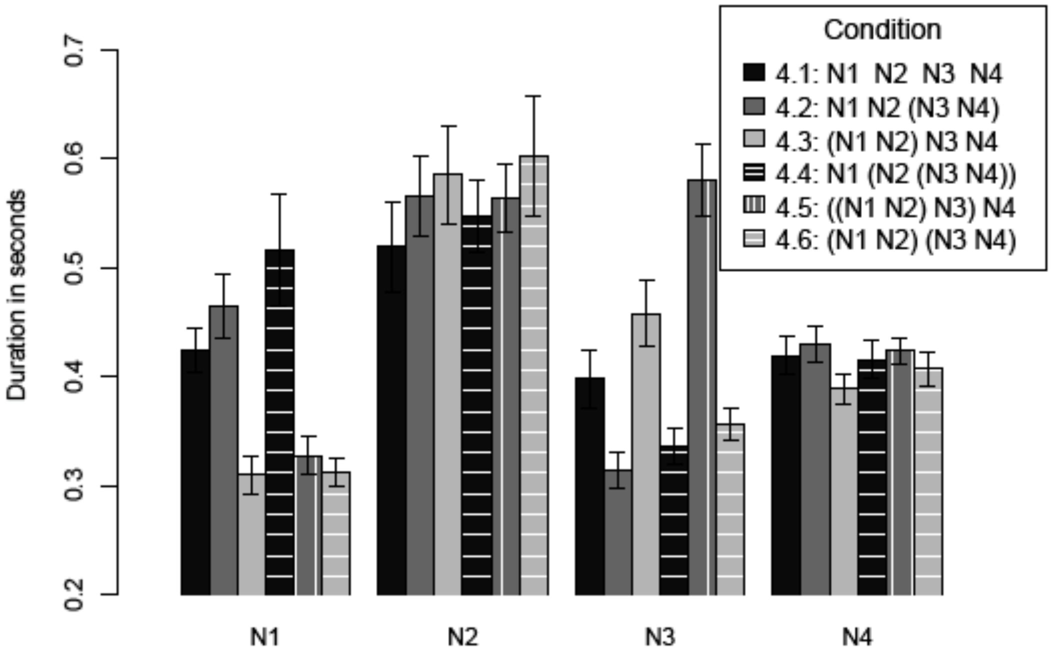
In the three-name conditions, pitch and duration deliver equivalent results in that higher pitch generally coincides with longer duration and lower pitch patterns with shorter duration. The pitch tracks reveal an interesting asymmetry: the right-branching pattern (3.3) has a striking resemblance to the baseline despite the difference on N2 – both have a downstep pattern. The left-branching pattern has a different shape, namely a clear upstep. But in the graph for duration, both patterns with groupings present clear differences in duration as compared to the baseline.

3.5 Results for four names

In this section, we compare the realizations 4.1 to 4.6. Figure 4 presents the results for duration for four names. The baseline (black) presents a short-long-short-long pattern, which is not accompanied by a similar effect in pitch (see below).

N2 of 4.3, 4.5 and 4.6 as well as N3 of 4.5 are all located in front of a (non-final) right parenthesis. These names are significantly lengthened compared to the baseline. And N1 of 4.4, which is in front of a left parenthesis, is also significantly lengthened. N1 of 4.2, a simplex element, is longer than the baseline but shorter than N1 of 4.4. And finally, N3 of 4.5, which is located in front of a right parenthesis, is much longer than all other third names. In contrast, N1 in the conditions 4.3, 4.5 and 4.6 is realized significantly shorter than the baseline; the same holds for N3 in 4.2, 4.4 and 4.6. These are all first elements of groupings. N2 of 4.4 is at the same time the first element of a recursive grouping and located before a left parenthesis. It has a similar duration to N2 of the baseline. Neutralization at the end of the sentence is once again observed in all patterns.

Figure 4: mean duration in ms of the four names per condition (German).



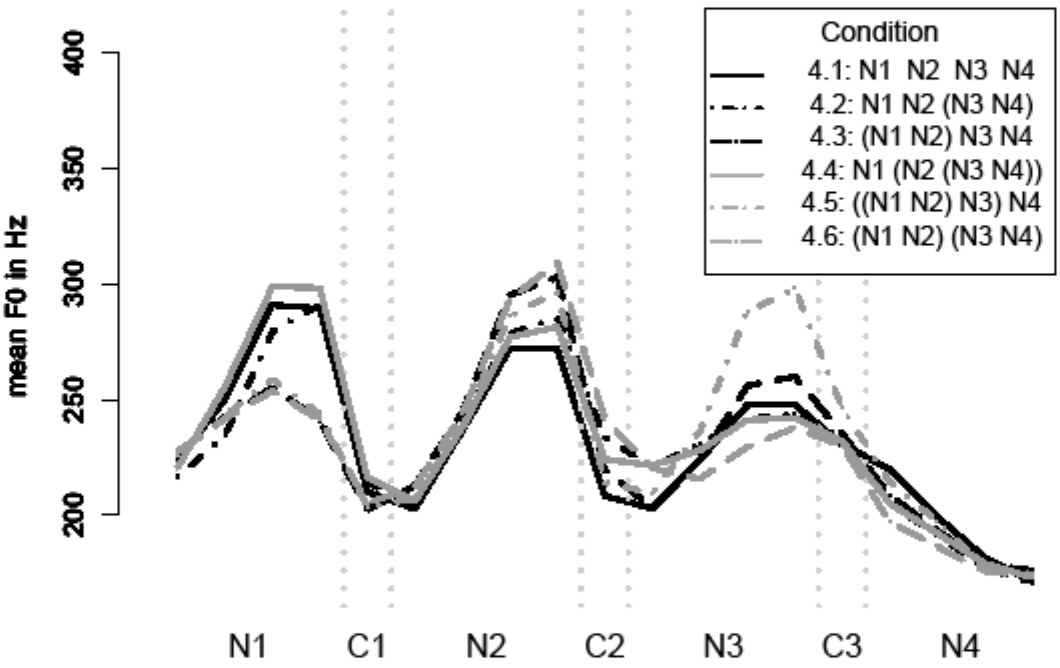
Summing up the results for duration in four names, the baseline shows the pattern short-long-short-neutralized. The effects of groupings are fine-grained and lead to a complex pattern of durational relationships.

Figure 5 shows the results for pitch. As was the case for the three-name patterns, the discussion concentrates on the relationship between the high tones of names. Except if it is the last one in the sentence, the rightmost element of a grouping is higher than the baseline. N2 in 4.3, 4.5 and 4.6 as well as N3 in 4.5 are the highest points in these sentences. In all three patterns, N1, the first element of the groupings, is then realized at a much lower level. N2 does not present very large differences in the absolute values, but in this half of the conditions (4.3, 4.5 and 4.6), an upstep from N1 to N2 can be observed, whereas in the other half, the value of N2 is the result of downstep relative to N1. This is the case in 4.1, 4.2 and 4.4, the structures without a right parenthesis after N2. These two conditions are the right-branching structures, and they have downstep throughout, thus resembling the baseline. However, in 4.2, 4.4 and 4.6, the third name is even lower than in the baseline, and, as was observed in the three-name patterns, the downstep between N2 and N3 is also larger than in the baseline, due to the fact that N3 is the first element of a grouping in these conditions. In 4.2 and 4.4, the elements before a parenthesis are higher than in the baseline. But the general impression is that 4.2 and 4.4 have the same shape as the baseline, and are subject to downstep throughout.

Returning to the left-branching structures, both 4.3 and 4.6 show a downstepped N3. The most interesting effect on N3 is the one in 4.5. Since it is the right-hand element of a grouping, it is strengthened. Its value is as high as N2, also the right-hand element of a grouping.

As before, the final values of all sentences are neutralized to a common value.

Figure 5: mean pitch in Hz of the four names per condition (German).



In sum, as already observed in the three-name pattern, we find parallel patterns of duration and pitch data in the four-name groupings, as well. A difference between pitch and duration concerns the baseline 4.1, which was shown to display a short-long-short-neutralized pattern in duration. As to the pitch contour, the baseline showed downstep throughout, and this was paralleled in the right-branching structures 4.2 and 4.4, though with interesting differences in height. The left-branching structures 4.3 and 4.5 show upstep as well as lengthening of the elements immediately preceding non-final right parentheses, again in different proportions. The first element of a grouping is lower in pitch and shorter at the same time.

3.6 Discussion

In the discussion, we compare our approach to the LRB and the SBR. We show that Proximity and Similarity are the best predictors for the distinctive results obtained for German.

The discussion is restricted to the four-name patterns, since the three-name patterns can for the most part be subsumed under them.

The baseline for four names 4.1 shows a rhythmic 2 x 2 pattern in duration. This imbalance is not paralleled in pitch, which shows regular downstep. Although small, this significant difference in duration is unexpected against the background of our assumptions. One possible explanation is the general tendency to shorten phrase initial elements, a phenomenon termed anacrusis (cf. Cruttenden 1997). Alternatively, there might be a preference for a binary short-long rhythm even in the absence of syntactic grouping.

In 4.2 and 4.3, the patterns with only one grouping, Wagner's (2005) SBR does not predict any difference in strength between the two simplex elements, nor any difference between them and the right-hand element of the grouping. All of them should be stronger than the baseline. Watson and Gibson's (2004) LRB, by contrast, predicts an asymmetry between the simplex elements: the one near the parenthesis should be stronger than the more distant one, which should have the same strength as the baseline. The results show that both simplex elements and the right-hand element of the groupings are stronger than the baseline, in line with the predictions of the SBR and Similarity. However the LRB and Anti-Proximity rightly predict an asymmetry in the two simplex elements, which the SBR cannot account for. Moreover, the pitch contour shows downstep in 4.2, the right-branching structure, like in the baseline, but not in 4.3, the left-branching structure. This difference is not expressed in the duration, which shows a short-long-short-neutralized pattern in all three cases. In the present account, it makes a difference whether the boundary at N2 is a left or a right parenthesis. N2 is subject to Anti-Proximity in both 4.2 and 4.3 and thus strengthened. But in 4.3, N1 is subject to Proximity, which weakens the boundary after it, creating a very large upstep between N1 and N2. In line with our predictions, N3 in 4.2 is lower than in 4.3 because it is subject to Proximity, whereas in 4.3, the strengthening effect of Similarity is at play.

As for the patterns with embedded groupings 4.4 and 4.5, both the SBR and the LRB, as well as Similarity and Proximity, predict a mirror structure, in which the boundary decreases (in 4.4) or increases (in 4.5) in strength from N1 to N4. This is confirmed by the results, except that the duration of N1 in 4.4 is only marginally weaker than that of N2. Again, the pitch shows downstep throughout in the right-branching structure 4.4, whereas it shows upstep in the left-branching 4.5. N2 in 4.4 is subject to both Proximity, as a weak element of a grouping, and Anti-Proximity, by virtue of being in front of a left parenthesis. As a result, it is neutralized (the difference in pitch and duration between N2 in 4.4. and N2 in the baseline is non-significant).

In 4.6, finally, all approaches predict a stronger boundary between N2 and N3 than between N1 and N2 on the one hand and between N3 and N4 on the other. In our data, the boundary after N2 in 4.6 is marked with significantly higher pitch and longer duration compared to the baseline. We interpret this strong boundary as an effect of Anti-Proximity, which exerts its force on N2, which is sandwiched between two elements (N1 and N3) that are weakened by Proximity.

In addition to the general parallelism of the two prosodic dimensions, our experiment revealed several differences between pitch and duration. First, as already mentioned, the baseline for four names 4.1 shows a rhythmic 2 x 2 pattern in duration. This imbalance is not paralleled in pitch, which shows regular downstep. Second, we observed an interesting asymmetry between left- and right-branching structures concerning pitch: while right-branching structures display a typical downstep pattern, following the baseline closely, left-branching structures are marked by clear upstep at the rightmost left branch. The apparent resemblance between right-branching structures and the baseline is not paralleled in the duration data, where both left- and right-branching structures deviate from the baseline patterns. Therefore, it may be assumed that pitch provides important information for the interpretation of syntactic effects on prosody which is lost when only duration is taken into consideration.

The doubly embedded syntactic structure is reflected in prosodic structure: embedding of elements in complex constituents is expressed both by upstep and lengthening (rightmost elements of groupings) and by prosodic compression (H-lowering and shortening on the left-hand elements of groupings). Left-branching induces upstep, while right-branching structures show a regular downstep.

In sum, apart from the rhythmic effect in the baseline, which is unexpected in all the models, Similarity and Proximity can account for all effects found in this experiment. While many of the effects were also predicted by the SBR or LRB, we have shown that our approach outperforms these models in predicting effects not captured by these models. In addition, the model not only accounts for durational effects but also for the distinctive pitch data.

4 Hindi

4.1 Methods and material

The entire data set used in this experiment was collected in one experiment, run individually with 20 speakers. All were female students of the University of Delhi, India. They were native speakers of Hindi in their twenties, coming mostly from Delhi and surrounding states. They were paid for their participation. Contexts had been previously recorded in a speech recording lab at the University of Potsdam by a native speaker of Hindi. The experiment was carried out as a Powerpoint presentation in a quiet room at the University of Delhi, and the outputs of the participants were recorded on a DAT tape recorder. Participants saw a question and its answer written in devanaagri script on the screen. They simultaneously heard the question over headphones. The questions were spoken by a male voice. Participants were instructed to speak out the names displayed on the screen as a natural response to the question they heard. Once they answered the question without any hesitation, the next stimulus was presented. In case of hesitation, they were asked to repeat the answer. Presentation flow was controlled by the experimenter. Participants were allowed to take a few minutes break whenever they wanted. A set of instructions and two practice examples familiarized the subjects with the process. Examples of the material, transposed into roman script, are shown in (15):

(12) Hindi examples

a. Context: Nihaar and Rohan always go together on picnics. Vinay is also interested in outings.
Question: With whom would you like to go for a picnic?

Answer: (nihaar aur rohana) yaa vinay ke saatha
(With Nihaar and Rohan) or Vinay

b. Context: Viral and Vaaman always practice singing together, just like Yaman and Yogi do.
Question: With whom would you like to practice singing?

Answer: (Viral aur vaaman) yaa (yaman aur yogi) ke saatha
(With Viral and Vaaman) or (Yaman and Yogi)

Altogether, 600 sentences entered the analysis, 120 with three names (20 subjects x 3 conditions x 2 contexts), and 480 sentences with four names (20 subjects x 6 conditions x 4 contexts). The smaller number of contexts for the three names, as compared to German, was the consequence of a mistake in the experimental procedure. Only two contexts were recorded for the baseline conditions. As a result, we decided to analyze only those two contexts in the other conditions as well, in order to have comparable data sets.[3]

It has to be noticed that the Hindi sentences differ from the German data in having a final postposition *ke saatha* 'with' as illustrated in (12). The last name is thus not sentence-final as it was in German. Thus it comes as even more of a surprise that the last names were neutralized in Hindi, as they were in German, as shown below. A further difference from German comes from the fact that more names were used with differing numbers of syllables. We therefore made sure that the names were counterbalanced across conditions. The names are listed in the Appendix.

Measurements and hypotheses

The same measurements and statistical analyses were conducted with the Hindi data as have been described above for the German material.

As for the hypotheses, the same ones as for German were assumed for Hindi as well.

4.2 Three names

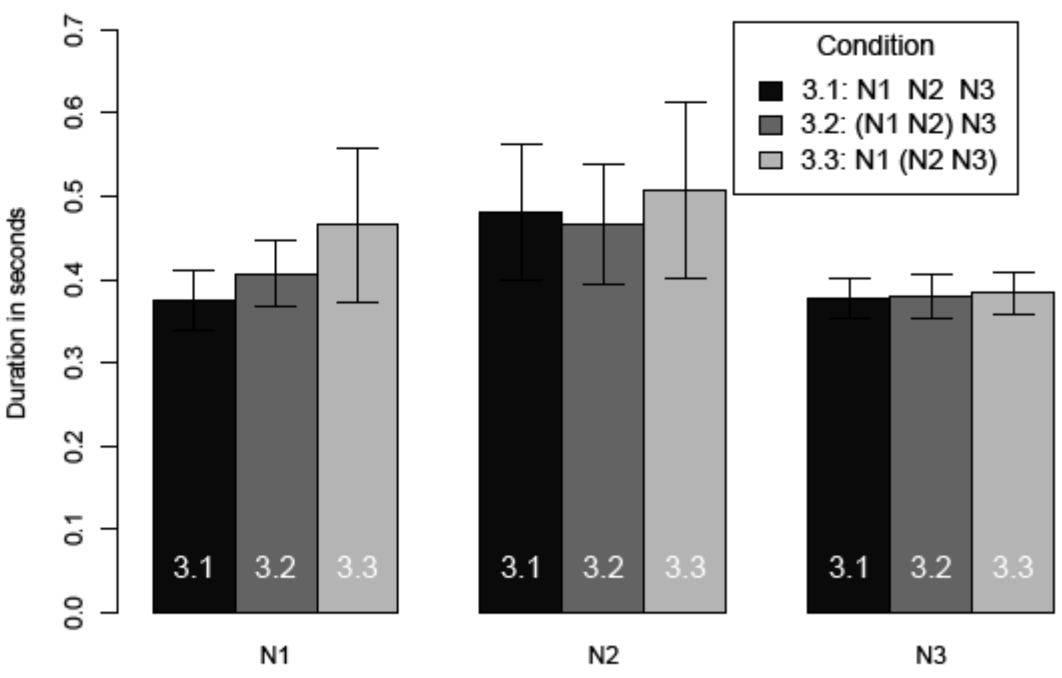
As we did for German, we first compare the realizations 3.2 and 3.3 with the baseline realization 3.1 without brackets.

First, let us examine the results for duration in Figure 6. The general impression is very different from what we observed for German. Only N1 shows variation between conditions. In the baseline pattern 3.1, N2 is longer than N1, and N3 is the shortest, as was observed for German.

[3] The results for all four contexts in the two non-baseline conditions are nearly indistinguishable from the results of the two analyzed contexts.

Condition 3.2 partly shows the same tendency as in German: N1 is significantly shorter than N2. However, N1 in 3.2 is slightly longer than in the baseline condition. This result goes against Proximity. Both in 3.2 and in 3.3 N2 and N3 have the same duration as in condition 3.1. In other words, no lengthening takes place at the end of the grouping in 3.2. In condition 3.3, N1 is longer than in the other conditions. It has the same length as N2. A name before a left parenthesis is longer than the baseline. But no shortening of N2 takes place at the beginning of the grouping. Proximity is not confirmed here.

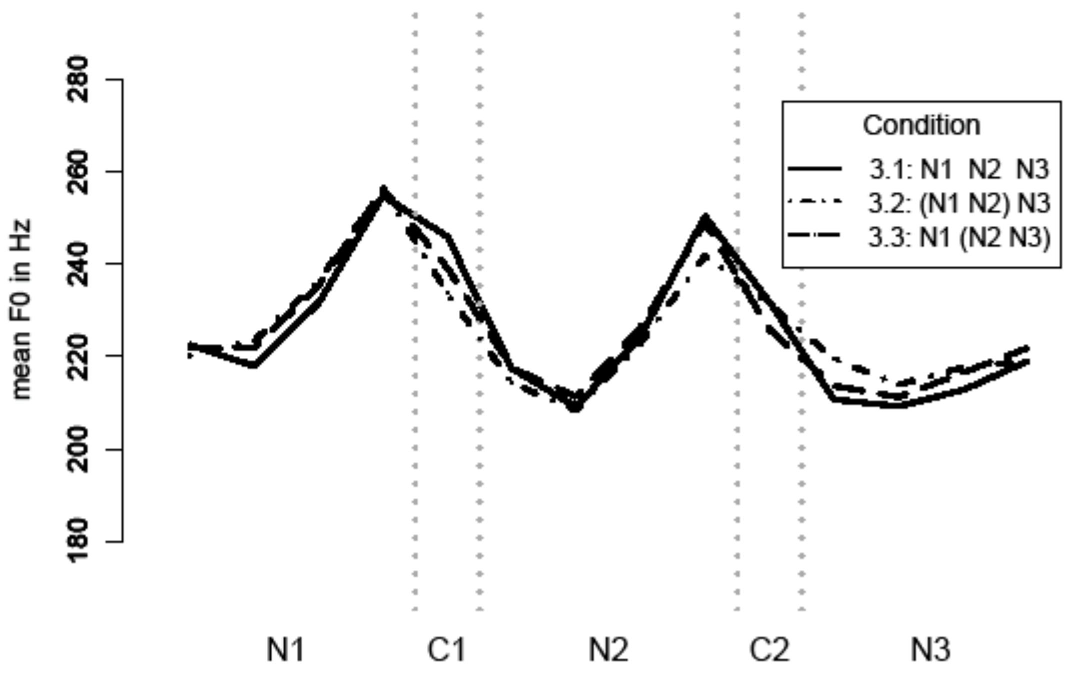
Figure 6: mean duration in ms of the three names per condition (Hindi).



To sum up the results for duration in the three-name conditions, the effect of grouping is restricted to N1 when N1 is a simplex element before a grouping (condition 3.3). In this case, it is lengthened compared to the baseline. N2 and N3 are unaffected.

Turning now to the results for pitch in the three-name realizations in Figure 7, the lack of effect of grouping is even more apparent. The differences that we could identify in the German contours are not visible in Hindi. The three normalized pitch tracks are more or less isomorphic in all three conditions. It can be seen that the highest point of the names is aligned further to the right in the three names than was the case for German. This is due to the tonal pattern of Hindi (see Patil et al. 2008). The high tone is the boundary tone of a p-phrase, located as far to the right in the p-phrase as possible.

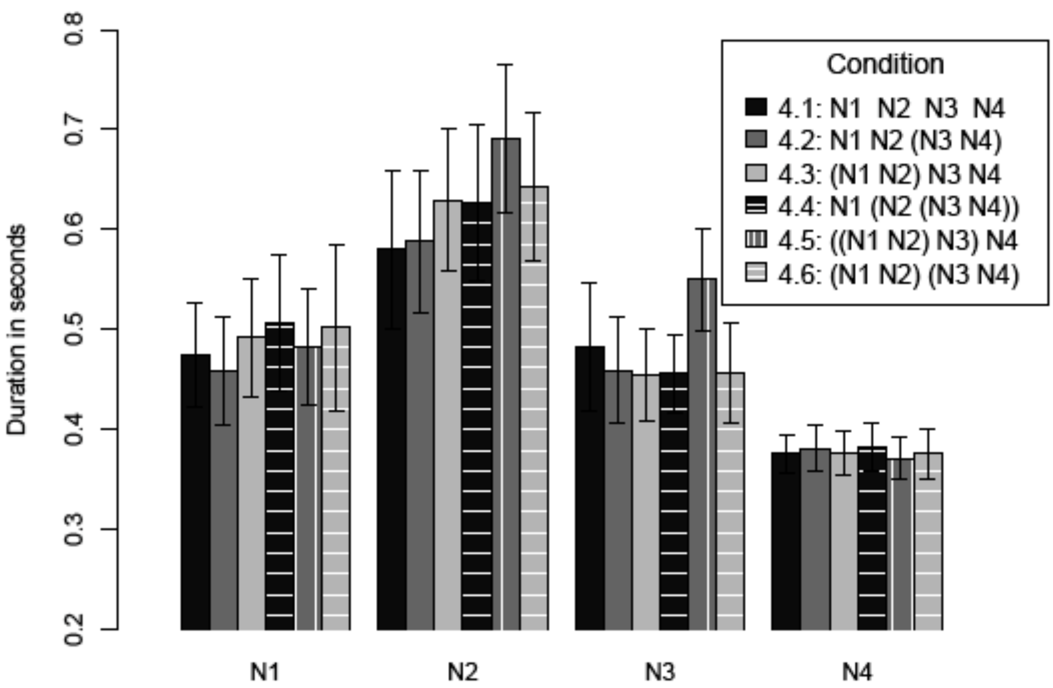
Figure 7: mean pitch in Hz of the three names per condition (Hindi).



4.3 Four names

The results for duration in the four-name conditions are displayed in Figure 8. As was the case for three names, all contours have the same general shape, though some differences in duration between the individual conditions are apparent. Comparing N1, no significant differences can be observed. N2 in condition 4.5 is longer than in the other conditions. N3 is also the longest in 4.5, all other conditions being slightly shorter than in the baseline. N4 is neutralized for its length, as was the case with N3 in the three-name conditions.

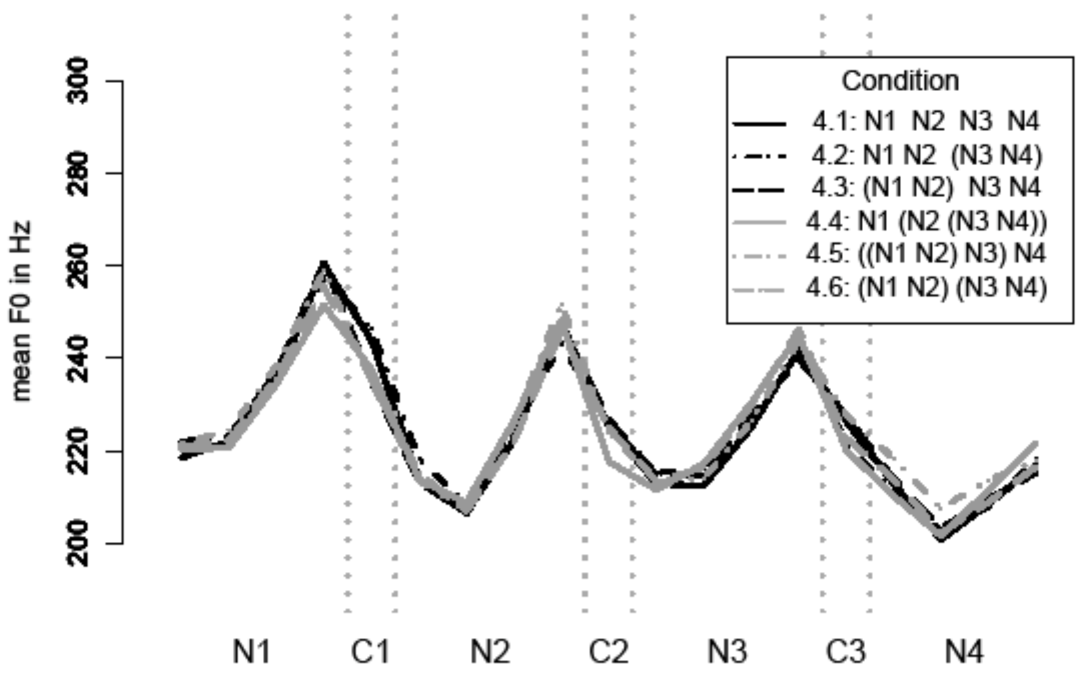
Figure 8: mean duration in ms of the four names per condition (Hindi).



The results for duration do not confirm our hypotheses. Except for the longer duration of N2 and N3 in 4.5, only minimal changes relative to the baseline are observed in the other conditions.

Figure 9 illustrates the pitch contours of the four-name sentences. Again, the absence of differences between the conditions is remarkable. All 6 conditions have exactly the same shape, and the contours are isomorphic.

Figure 9: mean pitch in Hz of the four names per condition (Hindi).



Again, the hypotheses formulated in section 4.2 are not confirmed. Hindi speakers do not contrast the different groupings of names by pitch.

4.4 Discussion

The data for Hindi show a great insensitivity to groupings as compared to those for German. We find a surprising lack of difference in the prosodic correlates between the conditions, and this both in the three- and in the four-name conditions. As for duration, the difference between N1 and N2 in the baseline could possibly be attributed to phrase-initial anacrusis (Cruttenden 1997).

The pitch contours are identical in all cases, and the graphs for duration show some differences, but not in the direction expected if prosody reflects syntax and recursion in the way predicted by Proximity and Similarity. In fact, the only data that is in line with our assumptions is the duration of N2 and N3 in condition 4.5, which is lengthened compared to the baseline. But this increase in duration is also compatible with an increase in processing difficulties.

5 General discussion

5.1 Interpretation of the experimental results

The results for German show that prosodic structure reflects grouping and embedding in a precise way, both for duration and for pitch. Proximity and Similarity account for the prosodic structure emerging from the syntactic structure. The first principle, Proximity, accounts for the lower pitch and shorter duration observed on the left-hand member of groupings. Its corollary, Anti-Proximity, has the opposite effect and strengthens the boundary between two constituents. The second principle, Similarity, accounts for the observation that simplex elements in an expression containing groupings have increased duration and higher pitch to achieve similar prosody to complex elements at the same level of syntactic embedding. As a result, German uses prosody in a syntax-sensitive way, interpreting syntactic structure in careful detail. We claim that this property of German correlates with its general intonational system. German, as an intonation language, is able to change pitch accents and boundary tones in a variety of ways to express pragmatic meanings. Pitch scaling is a fine-grained device which supports this use of intonation, as demonstrated by Féry & Kügler (2008). Our experiment demonstrates that prosody as a whole is a support for the rendition of syntactic structure.

Hindi, by contrast, shows a surprising lack of correlation between syntactic structure and prosody. Neither Proximity nor Similarity were supported by the Hindi data. The SBR and the LRB are also unable to make the correct predictions for Hindi. These results can only be understood when Hindi intonation is considered as a whole. Hindi is a phrase language, according to the sentence-based typology of intonational systems. The melody of a sentence arises not because of pitch accents but primarily because of the distribution of phrasal tones, which are, as the name of these tones indicates, assigned at the level of the phrase. High tones in a syntactically simple Hindi sentence are always in a downstep relation, and are only marginally sensitive to information structure (see Patil et al. 2008).

Proximity is forward oriented. The realization of a constituent always modulates the distance to the following constituent, and not that to the preceding one (see also Kentner 2007 for a similar conclusion). The question may be asked whether Hindi has the reverse strategy, and thus reduces or increases the distance to the preceding constituent. Nothing indicates that this is correct. There is no prosodic change whatsoever manipulating the relationship of an element either to its left or to its right.

5.2 Implications

The results of this double experiment have some implications for theoretical considerations about the role of prosody in reflecting syntax. Clearly it seems that the prosodic reflection of syntactic and/or semantic structure is not a universal property as has been suggested by Yang (2007) for the parameter of duration. Instead, it crucially depends on the language-specific role of prosody whether and to what extent syntactic relations are reflected in prosodic structure. The apparent lack of prosodic marking of syntactic grouping in Hindi might be the cost for the clear and consistent marking of prosodic phrases, which do not adjust to pragmatic conditions in the same way as in German. This difference is important for understanding the role that prosody plays in language comprehension. It might be that languages may differ in this dimension much more than assumed thus far. The research on the role of prosody in speech processing has largely concentrated on intonation languages, which use pitch changes and pitch scaling for the communication of syntax and semantics, and has largely ignored the phrase languages, which rely more on phrasing for this parameter. We hope to have revealed the need for well-designed experiments to shed more light on this issue.

6 Conclusion

The results of two production experiments on name grouping, one in German and one in Hindi, have shown that these two languages differ as to how they involve prosody in syntactic disambiguation. Whereas German changes duration and pitch in a fine-grained manner, displaying a different prosodic pattern for each of the nine structures investigated, Hindi shows a surprising lack of these grammatical devices for the same task. We discuss the results of the two experiments in detail, and present a new analysis of the produced sentences for German. Two principles are shown to explain the data: Proximity, which translates the syntactic relationship between a constituent and the following one into prosodic phrasing, and Similarity, which adjusts the prominence level of simplex constituents to those of complex constituents at the same level of embedding. We show that Similarity and Proximity make better predictions than the LRB and the SBR, at least for German.

Hindi does not conform to these principles: the prosodic structure of our data does not show much sensitivity to syntactic structure, and the few changes obtained may be due to a general increase in processing difficulty. We propose that the obvious difference between the two languages correlates with their different intonation systems. German is an intonation language, using distinctions in pitch accents to express differences in pragmatic meanings, and Hindi is a phrase language. In this kind of language, tonal events are mostly correlated with phrasal tones rather than with pitch accents.

In short, prosody is not used in the same way in all languages to express syntactic structures, and it is an important task for the future to understand how exactly languages differ on this issue.

7 References

Boersma, P. and D. Weenink (2008). Praat: doing phonetics by computer [Computer program]. Retrieved from: http://www.praat.org.

Carlson, K., C. Clifton, and L. Frazier (2001). Prosodic boundaries in adjunct attachment. Journal of Memory and Language 45, 58–81.

Clifton, C., K. Carlson, and L. Frazier (2002). Informative prosodic boundaries. Language and Speech, 45, 87-114.

Cooper, W.E. and J. Paccia-Cooper (1980). Syntax and Speech. Harvard University Press.

Cruttenden, A. (1997). Intonation. Cambridge University Press.

Ferreira, F. (1993). Creation of prosody during sentence production. Psychological review, 100 (2), 233-253.

Féry, C. (2008). Indian languages as intonational ‘phrase languages’. To appear in a secret Festschrift.

Féry, C. and F. Kügler (2008). Pitch accent scaling in German. Journal of Phonetics. (still to appear)

Féry, C. & V. Samek-Lodovici (2006). Focus projection and prominence in nested foci. Language, 82, 131-150.

Gee, J. and F. Grosjean (1983). Performance structures: A psycholinguistic and linguistic appraisal. Cognitive Psychology 15, 411-458.

Grabe, E. (1998). Pitch accent realization in English and German. Journal of Phonetics 26, 129-143.

Hunyadi, L. (2006). Grouping, the cognitive basis for recursion in language. Argumentum 2, 67-114.

Kentner, G. (2007). Length, ordering preferences and intonational phrasing: Evidence from pauses. Interspeech-2007, 2337-2340.

Lehiste, I. (1983). The many linguistic functions of duration. Zürich: Symposium Prosody.

Levelt, W.J.M. (1989). Speaking: From intention to articulation. Cambridge MA: MIT Press.

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

Patil, U., G. Kentner, A. Gollrad, F. Kügler, C. Féry and S. Vasishth (2008). Focus, word order and intonation in Hindi. Journal of South Asian Linguistics.

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. Dissertation. MIT.

Phillips, C. (2003). Linear order and constituency. Linguistic Inquiry 34, 37–90.

Raczaszek, J. B. Tuller, L.P. Shapiro, P. Case, and S. Kelso 1999. Categorization of ambiguous sentences as a function of a changing prosodic parameter: A dynamical approach. J. Psycholinguistic Research 28, 367-393.

Selkirk, E. (1984). Phonology and syntax: The relationship between sound and structure. MIT Press, Cambridge, MA.

Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. Goldsmith (Ed.). The Handbook of Phonological Theory, 550–569.

Truckenbrodt, H. (2002). Upstep and embedded register levels. Phonology 19, 77-120.

Wagner, M. 2005. Prosody and recursion. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Watson, D. and E. Gibson (2004). The relationship between intonational phrasing and syntactic structure in language production. Language and Cognitive Processes 19/6, 713-755.

Wightman, C.W., S. Shattuck-Hufnagel, M. Ostendorf and P.J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. Journal of the Acoustical Society of America, 92, 1707-1717.

Yang, Li-chiung (2007). Duration and pauses as boundary markers in speech: A cross-linguistic study, In INTERSPEECH-2007, 458-461.

Appendix:

I. Stimuli

A: German experiment

List of names in German experiment: Rudi, Anja, Lina, Suse, Willi, Nino, Erna, Jana, Anna, Mila, Nina, Maya, Arno
Conjunctions: oder – *or*; und – *and*

Complete list of items:

3.1 01 Rudi oder Anja oder Lina.

3.1 02 Rudi oder Erna oder Nino.

3.1 03 Anna oder Suse oder Jana.

3.1 04 Willi oder Nino oder Anna.

3.2 01 (Anja und Mila) oder Jana.

3.2 02 (Lina und Suse) oder Anja.

3.2 03 (Suse und Lena) oder Willi.

3.2 04 (Mila und Erna) oder Nino.

3.3 01 Nino oder (Mila und Erna).

3.3 02 Willi oder (Suse und Lena).

3.3 03 Jana oder (Anja und Mila).

3.3 04 Anja oder (Lina und Suse).

4.1 01 Willi oder Nino oder Anna oder Mila.

4.1 02 Anja oder Erna oder Rudi oder Mila.

4.1 03 Suse oder Lina oder Willi oder Nino.

4.1 04 Anna oder Lina oder Rudi oder Mila.

4.2 01 Jana oder Anna oder (Mila und Erna).

4.2 02 Rudi oder Lina oder (Suse und Lena).

4.2 03 Jana oder Anna oder (Anja und Mila).

4.2 04 Nino oder Rudi oder (Lina und Suse).

4.3 01 (Lina und Suse) oder Rudi oder Nino.

4.3 02 (Anja und Mila) oder Jana oder Anna.

4.3 03 (Mila und Erna) oder Jana oder Anna.

4.3 04 (Suse und Lena) oder Rudi oder Lina.

4.4 01 Anna oder (Mila und (Lina und Suse)).

4.4 02 Suse oder (Willi und (Anja und Mila)).

4.4 03 Lina oder (Maya und (Mila und Erna)).

4.4 04 Nino oder (Rudi und (Suse und Lina)).

4.5 01 ((Suse und Lena) und Rudi) oder Nino.

4.5 02 ((Mila und Erna) und Maya) oder Lina.

4.5 03 ((Lina und Suse) und Mila) oder Anna.

4.5 04 ((Anja und Mila) und Willi) oder Suse.

4.6 01 (Susi und Lena) oder (Maya und Arno).

4.6 02 (Anja und Mila) oder (Jana und Nino).

4.6 03 (Mila und Erna) oder (Rudi und Anna).

4.6 04 (Lina und Suse) oder (Jana und Anja).

B: Hindi experiment

List of names in Hindi experiment: Raam, Aman, Bimal, Karan, Aramaan, Mohan, Nil, Vibhav, Nihaar, Rohan, Vinay, Raaghav, Viral, Vaaman, Yaman, Yogi
Conjunctions: yaa – *or*; aur – *and*

Complete list of items:

3.1 01 raam yaa aman yaa bimal ke saatha

3.1 03 nihaar yaa rohan yaa vinay ke saatha

3.2 01 (raam aur aman) yaa bimal ke saatha

3.2 03 (nihaar aur rohan) yaa vinay ke saatha

3.3 01 raam yaa (amana yaa bimal) ke saatha

3.3 03 nihaar yaa (rohana aur vinay) ke saatha

4.1 01 raam yaa aman yaa bimal yaa karaN ke saatha

4.1 02 aramaan yaa mohan yaa nIl yaa vibhav ke saatha

4.1 03 nihaar yaa rohan yaa vinay yaa raaghav ke saatha

4.1 04 viral yaa vaaman yaa yaman yaa yogI ke saatha

4.2 01 raam yaa aman yaa (bimal aur karaNa) ke saatha

4.2 02 aramaan yaa mohan yaa (nIl aur vibhava) ke saatha

4.2 03 nihaar yaa rohan yaa (vinay aur raaghava) ke saatha

4.2 04 viral yaa vaaman yaa (yaman aur yogI) ke saatha

4.3 01 (raam aur amana) yaa bimal yaa karaN ke saatha

4.3 02 (aramaan aur mohana) yaa nIl yaa vibhav ke saatha

4.3 03 (nihaar aur rohana) yaa vinay yaa raaghav ke saatha

4.3 04 (viral aur vaamana) yaa yaman yaa yogI ke saatha

4.4 01 raam yaa (aman aur (bimal aur karaNa)) ke saatha

4.4 02 aramaan yaa (mohan aur (nIl aur vibhava)) ke saatha

4.4 03 nihaar yaa (rohan aur (vinay aur raaghava)) ke saatha

4.4 04 viral yaa (vaaman aur (yaman aur yogI)) ke saatha

4.5 01 ((raam aur amana) aur bimala) yaa karaN ke saatha

4.5 02 ((aramaan aur mohana) aur nIla) yaa vibhav ke saatha

4.5 03 ((nihaar aur rohana) aur vinaya) yaa raaghav ke saatha

4.5 04 ((viral aur vaamana) aur yamana) yaa yogI ke saatha

4.6 01 (raam aur amana) yaa (bimal aur karaNa) ke saatha

4.6 02 (aramaan aur mohana) yaa (nIl aur vibhava) ke saatha

4.6 03 (nihaar aur rohana) yaa (vinay aur raaghava) ke saatha

4.6 04 (viral aur vaamana) yaa (yaman aur yogI) ke saatha

II. Results

A: results for German experiment

1. Pairwise comparisons of the 3 conditions for German 3 Names experiment: Mean differences and 95% confidence intervals for duration (dur in ms) and maximum F0 (F0max in Hertz) for first (N1) and second name (N2) respectively:

pairwise comparisons N1			Cond. 3.2 dur	Cond. 3.2 F0max	Cond. 3.3 dur	Cond. 3.3 F0max
Cond. 3.1	mean difference		109	46	-86	-5
	95% CI	lower	79	38	-115	-14
		upper	138	55	-56	3
	p-value		<0.001	<0.001	<0.001	>0.1
Cond. 3.2	mean difference				-195	-52
	95% CI	lower			-225	-60
		upper			-165	-43
	p-value				<0.001	<0.001

pairwise comparisons N2			Cond. 3.2 dur	Cond. 3.2 F0max	Cond. 3.3 dur	Cond. 3.3 F0max
Cond. 3.1	mean difference		-63	-43	137	12
	95% CI	lower	-88	-51	111	4
		upper	-37	-35	163	20
	p-value		<0.001	<0.001	<0.001	0.001
Cond. 3.2	mean difference				200	55
	95% CI	lower			174	47
		upper			225	63
	p-value				<0.001	<0.001

2. Pairwise comparisons of the 6 conditions for German 4 Names experiment: Mean differences and 95% confidence intervals for duration (dur in ms) and maximum F0 (F0max in Hertz) for first (N1), second (N2), and third name (N3) respectively:

pairwise comparisons N1			Cond. 4.2		Cond. 4.3		Cond. 4.4		Cond. 4.5		Cond. 4.6	
			dur	F0max	dur	F0max	dur	F0max	dur	F0max	dur	F0max
4.1	mean difference		-40	1	115	42	-88	-12	98	38	113	42
	95% CI	lower	-75	-11	80	30	-123	-24	62	26	78	30
		upper	-5	13	150	54	-53	0	132	50	148	54
	p-value		0.013	>0.1	0.001	<0.001	0.001	0.064	0.001	<0.001	0.001	<0.001
4.2	mean difference				155	41	-48	-13	137	37	153	41
	95% CI	lower			120	29	-82	-25	102	24	118	29
		upper			197	53	-13	-1	172	49	188	53
	p-value				<0.001	<0.001	0.001	0.025	<0.001	<0.001	<0.001	<0.001
4.3	mean difference						-202	-54	-18	-4	-2	0
	95% CI	lower					-237	-66	-53	-16	-37	-12
		upper					-167	-41	18	8	33	12
	p-value						<0.001	<0.001	>0.1	>0.1	>0.1	>0.1
4.4	mean difference								185	50	201	54
	95% CI	lower							150	38	166	42
		upper							220	62	236	66
	p-value								<0.001	<0.001	<0.001	<0.001
4.5	mean difference										16	4
	95% CI	lower									-19	-8
		upper									51	17
	p-value										>0.1	>0.1

pairwise comparisons N2			Cond. 4.2		Cond. 4.3		Cond. 4.4		Cond. 4.5		Cond. 4.6	
			dur	F0max	dur	F0max	dur	F0max	dur	F0max	dur	F0max
4.1	mean difference		-46	-13	-65	-30	-26	-7	-44	-24	-82	-34
	95% CI	lower	-96	-24	-115	-42	-76	-20	-94	-35	-132	-47
		upper	4	1	-15	-17	24	4	6	-11	-32	-22
	p-value		0.105	0.032	0.002	<.001	>0.1	>0.1	>0.1	<.001	<0.001	<.001
4.2	mean difference				-20	-17	20	5	2	-11	-36	-22
	95% CI	lower			-70	-29	-30	-7	-48	-23	-86	-34
		upper			31	-5	70	17	52	1	14	-10
	p-value				>0.1	0.001	>0.1	>0.1	>0.1	>0.1	>0.1	<.001
4.3	mean difference						39	22	21	6	-17	-5
	95% CI	lower					-10	10	-29	-6	-11	-16
		upper					89	34	71	18	33	7
	p-value						>0.1	<.001	>0.1	>0.1	>0.1	>0.1
4.4	mean difference								18	-16	-37	-27
	95% CI	lower							-42	-28	-88	-39
		upper							68	-4	2	-15
	p-value								>0.1	0.002	>0.1	<0.001
4.5	mean difference										38	-11
	95% CI	lower									-12	-23
		upper									88	1
	p-value										>0.1	>0.1

pairwise comparisons N3			Cond. 4.2 dur	Cond. 4.2 F0max	Cond. 4.3 dur	Cond. 4.3 F0max	Cond. 4.4 dur	Cond. 4.4 F0max	Cond. 4.5 dur	Cond. 4.5 F0max	Cond. 4.6 dur	Cond. 4.6 F0max
4.1	mean difference		85	5	-60	-13	61	6	-182	-54	42	14
	95% CI	lower	49	-5	-95	-23	26	-5	-217	-65	7	4
		upper	120	16	-25	-2	97	16	-147	-44	77	24
	p-value		<0.001	>0.1	<0.001	0.005	<0.001	>0.1	<0.001	<0.001	0.008	0.001
4.2	mean difference				-144	-18	-23	0	-266	-60	-43	8
	95% CI	lower			-179	-28	-58	-10	-302	-70	-78	-2
		upper			-109	-8	12	10	-231	-50	-8	18
	p-value				<0.001	<0.001	>0.1	>0.1	<0.001	<0.001	0.006	>0.1
4.3	mean difference						121	18	-122	-42	101	26
	95% CI	lower					86	7	-157	-52	66	16
		upper					157	28	-87	-32	137	37
	p-value						<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
4.4	mean difference								244	-60	-20	8
	95% CI	lower							208	-70	-55	-2
		upper							279	-50	15	18
	p-value								<0.001	<0.001	>0.1	>0.1
4.5	mean difference										-224	68
	95% CI	lower									-259	58
		upper									-188	79
	p-value										<0.001	<0.001

B: results for Hindi experiment

1. Pairwise comparisons of the 3 conditions for Hindi 3 Names experiment: Mean differences and 95% confidence intervals for duration (dur in ms) and maximum F0 (F0max in Hertz) for first (N1) and second name (N2) respectively:

pairwise comparisons N1			Cond. 3.2		Cond. 3.3	
			dur	F0max	dur	F0max
Cond. 3.1	mean difference		-25	-4	-63	1
	95% CI	lower	-81	-13	-12	-9
		upper	31	5	-6	10
	p-value		>0.1	>0.1	0.025	>0.1
Cond. 3.2	mean difference				-37	5
	95% CI	lower			-94	10
		upper			19	14
	p-value				>0.1	>0.1

pairwise comparisons N2			Cond. 3.2		Cond. 3.3	
			dur	F0max	dur	F0max
Cond. 3.1	mean difference		-7	5	0.3	0.3
	95% CI	lower	-96	-6	-89	-11
		upper	82	17	89	12
	p-value		>0.1	>0.1	>0.1	>0.1
Cond. 3.2	mean difference				7	-5
	95% CI	lower			-83	-17
		upper			97	7
	p-value				>0.1	>0.1

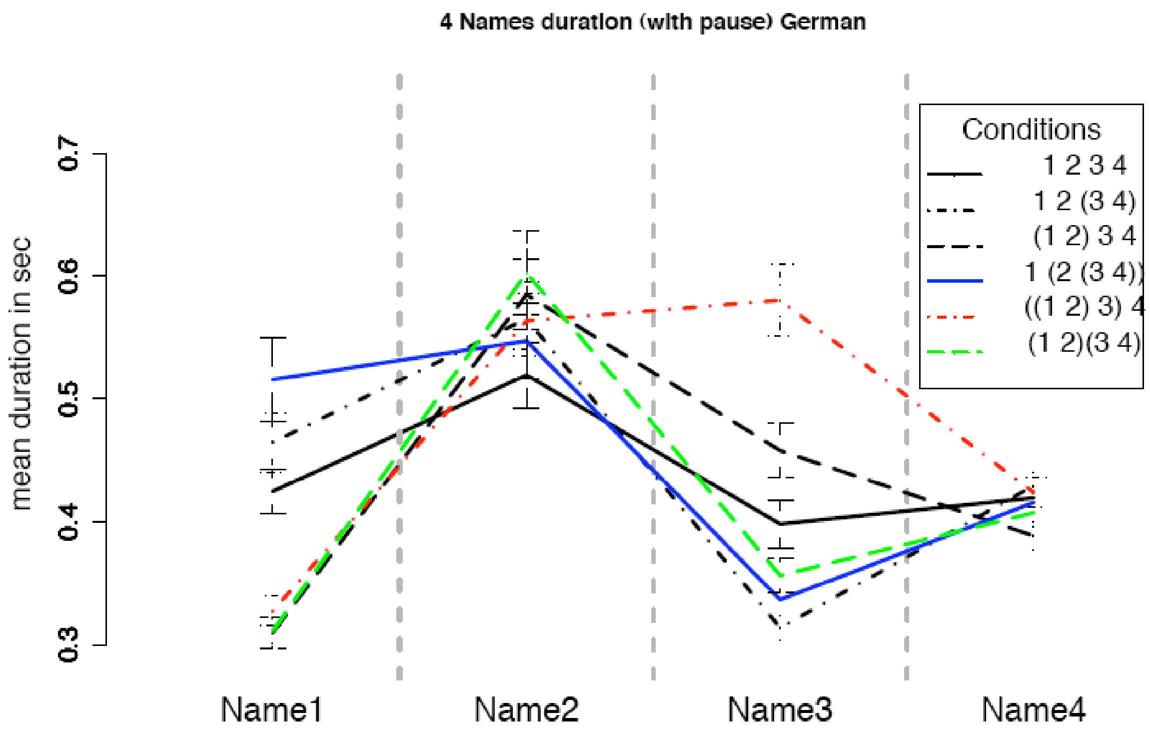
2. Pairwise comparisons of the 6 conditions for Hindi 4 Names experiment: Mean differences and 95% confidence intervals for duration (dur in ms) and maximum F0 (F0max in Hertz) for first (N1) and second (N2) and third name (N3) respectively:

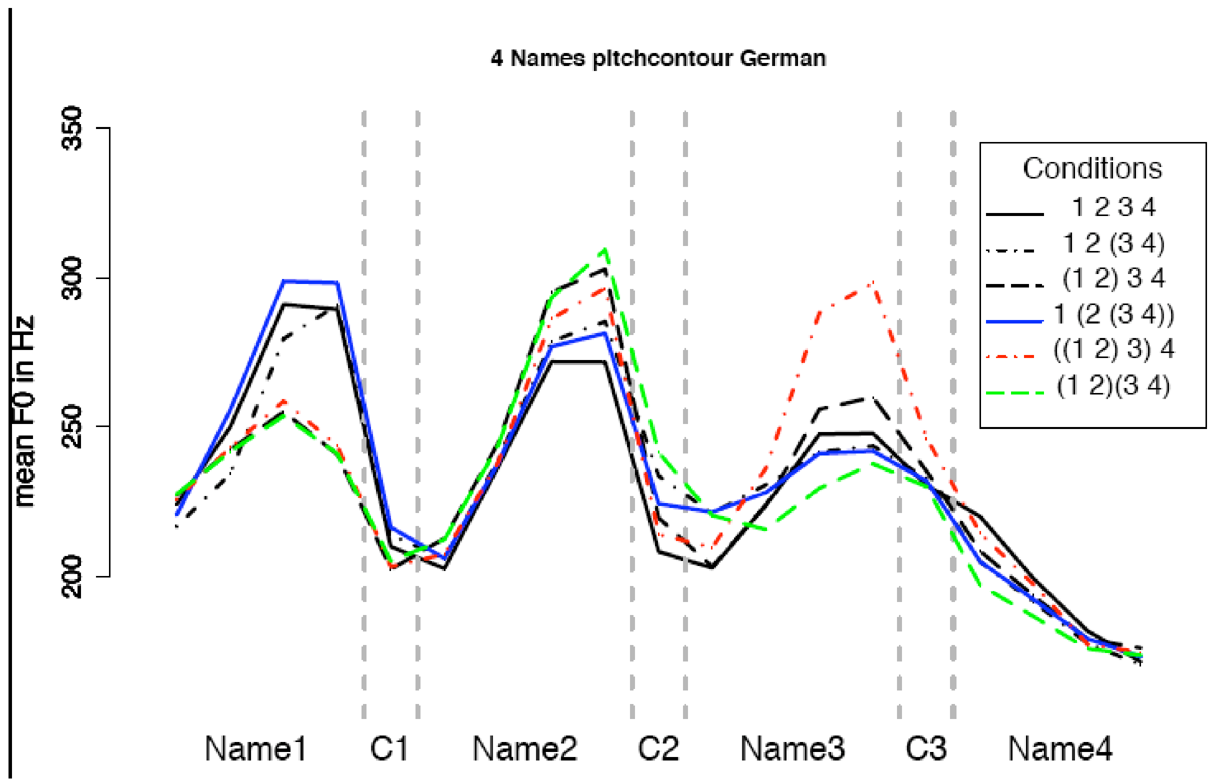
pairwise comparisons N1			Cond. 4.2		Cond. 4.3		Cond. 4.4		Cond. 4.5		Cond. 4.6	
			dur	F0max	dur	F0max	dur	F0max	dur	F0max	dur	F0max
4.1	mean difference		20	2	-17	-2	-31	5	-14	-1	-28	4
	95% CI	lower	-39	-6	-77	-10	-91	-3	-75	-9	-87	-5
		upper	80	10	42	6	28	13	48	7	32	12
	p-value		>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1
4.2	mean difference				-38	-4	-52	3	-35	-3	-48	2
	95% CI	lower			-98	-12	-112	-5	-96	-11	-108	-7
		upper			22	4	7	11	27	6	12	10
	p-value				>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	>0.1
4.3	mean difference						-14	7	4	1	-11	5
	95% CI	lower					-74	-2	-58	-8	-70	-3
		upper					46	14	65	9	50	13
	p-value						>0.1	>0.1	>0.1	>0.1	>0.1	>0.1
4.4	mean difference								18	-6	4	-1
	95% CI	lower							-44	-14	-56	-9
		upper							79	2	64	7
	p-value								>0.1	>0.1	>0.1	>0.1
4.5	mean difference										-14	5
	95% CI	Lower									-75	-4
		upper									48	13
	p-value										>0.1	>0.1

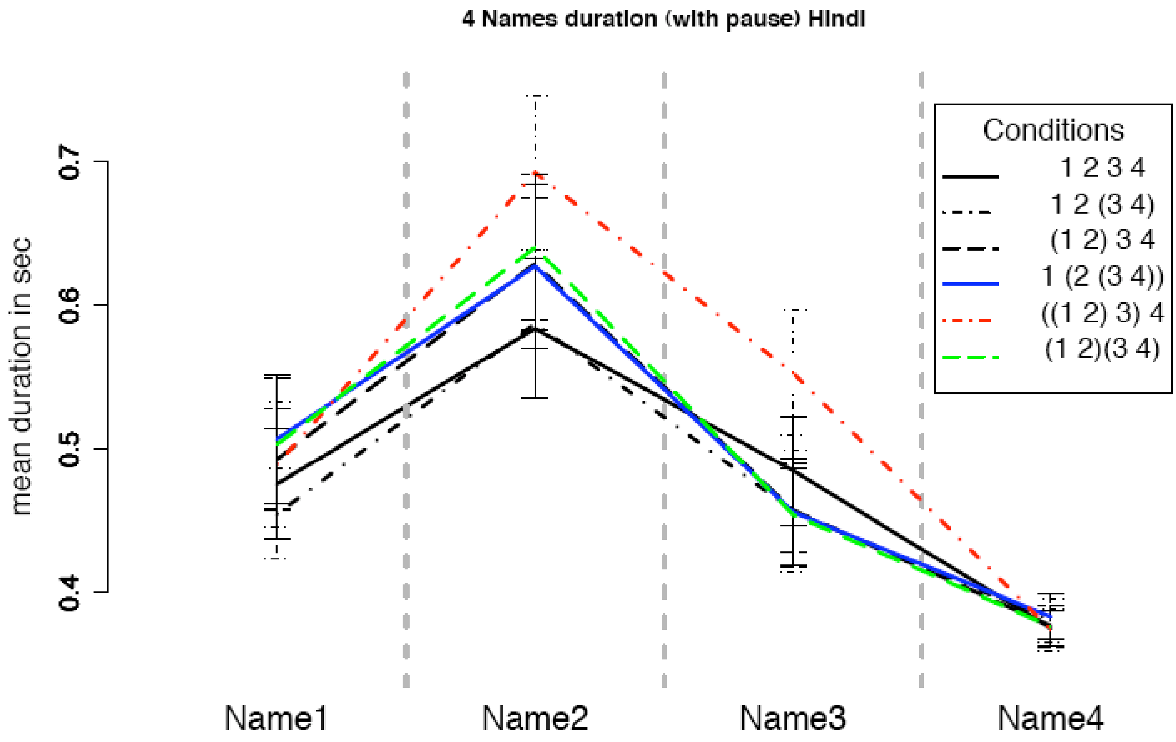
pairwise comparisons N2			Cond. 4.2		Cond. 4.3		Cond. 4.4		Cond. 4.5		Cond. 4.6	
			dur	F0max	dur	F0max	dur	F0max	dur	F0max	dur	F0max
4.1	mean difference		-3	3	-45	1	-43	-2	-108	-6	-56	2
	95% CI	Lower	-90	-5	-131	-7	-129	-10	-196	-14	-143	-7
		upper	83	12	41	10	43	6	-20	3	29	10
	p-value		>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	0.005	>0.1	>0.1	>0.1
4.2	mean difference				-42	-2	-40	-5	-105	-9	-53	-2
	95% CI	Lower			-128	-10	-126	-14	-194	-18	-139	-10
		Upper			45	6	46	3	-16	1	33	7
	p-value				>0.1	>0.1	>0.1	>0.1	0.005	>0.1	>0.1	>0.1
4.3	mean difference						-2	-3	-63	-7	-12	1
	95% CI	Lower					-84	-12	-152	-16	-98	-8
		Upper					88	5	25	1	75	9
	p-value						>0.1	>0.1	>0.1	>0.1	>0.1	>0.1
4.4	mean difference								-65	-4	-13	4
	95% CI	Lower							-153	-12	-100	-5
		Upper							23	5	73	12
	p-value								>0.1	>0.1	>0.1	>0.1
4.5	mean difference										52	8
	95% CI	Lower									-37	-1
		Upper									140	16
	p-value										>0.1	>0.1

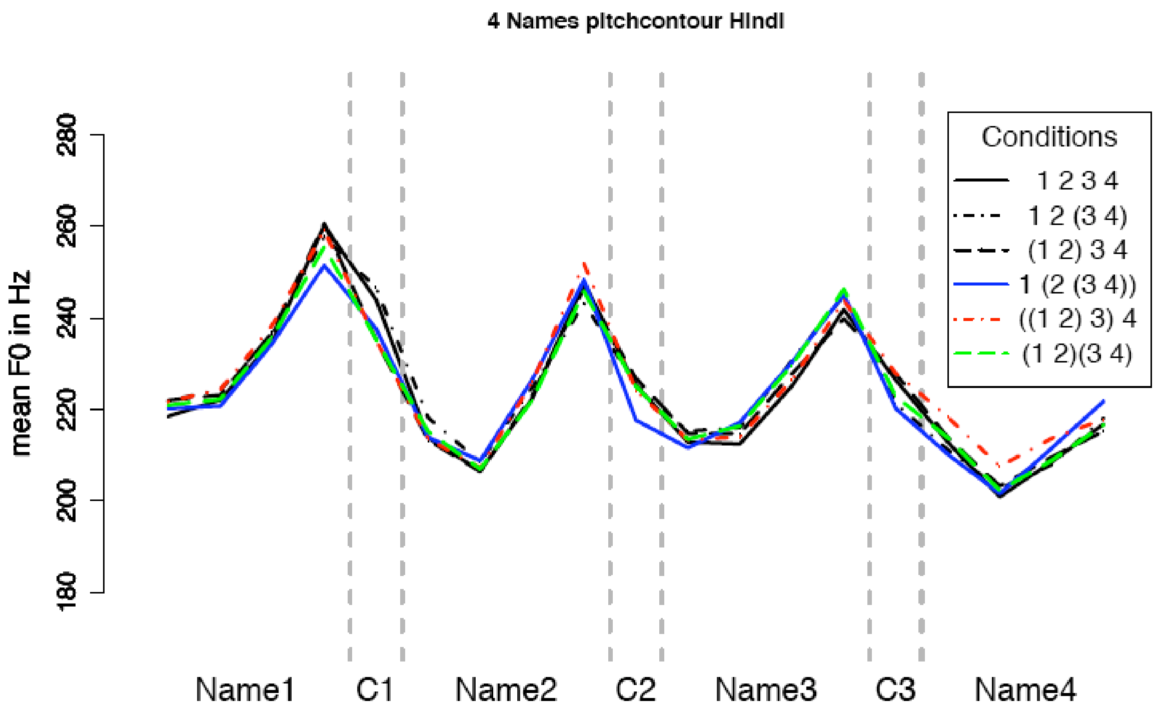
pairwise comparisons N3			Cond. 4.2		Cond. 4.3		Cond. 4.4		Cond. 4.5		Cond. 4.6	
			dur	F0max	dur	F0max	dur	F0max	dur	F0max	dur	F0max
4.1	mean difference		28	-3	27	-1	29	-4	68	-2	31	-4
	95% CI	lower	-39	-11	-39	-8	-37	-12	-136	-11	-36	-12
		upper	95	5	94	8	95	4	1	6	97	4
	p-value		>0.1	>0.1	>0.1	>0.1	>0.1	>0.1	0.056	>0.1	>0.1	>0.1
4.2	mean difference				0	3	1	-1	-96	1	3	1
	95% CI	lower			-67	-5	-65	-8	-165	-7	-64	-9
		upper			67	11	68	7	-27	9	70	7
	p-value				>0.1	>0.1	>0.1	>0.1	0.001	>0.1	>0.1	>0.1
4.3	mean difference						1	-4	-95	-2	3	-4
	95% CI	lower					-66	-12	-167	-11	-64	-12
		upper					68	5	-27	6	70	4
	p-value						>0.1	>0.1	0.001	>0.1	>0.1	>0.1
4.4	mean difference								-97	1	2	1
	95% CI	lower							-166	-7	-65	-8
		upper							-28	10	68	8
	p-value								0.001	>0.1	>0.1	>0.1
4.5	mean difference										99	-2
	95% CI	lower									30	-10
		upper									168	7
	p-value										<0.001	>0.1

Additional figures


4 Names duration (with pause) German
mean duration in sec
0.3
0.4
0.5
0.6
0.7
Conditions
1 2 3 4
1 2 (3 4)
(1 2) 3 4
1 (2 (3 4))
((1 2) 3) 4
(1 2)(3 4)
Name1
Name2
Name3
Name4


4 Names pitchcontour German
mean F0 in Hz
200
250
300
350
Conditions
1 2 3 4
1 2 (3 4)
(1 2) 3 4
1 (2 (3 4))
((1 2) 3) 4
(1 2)(3 4)
Name1
C1
Name2
C2
Name3
C3
Name4





Caroline Féry <caroline.fery@googlemail.com>
Gerrit Kentner gerrit.kentner@googlemail.com