Available online at www.sciencedirect.com

SciVerse ScienceDirect

Lingua 131 (2013) 29–48

Lingua

www.elsevier.com/locate/lingua

Disambiguating the scope of negation by prosodic cues in three varieties of German

Stefan Baumann [a,*], Tamara Rathcke [b]

[a] IfL Phonetik, University of Cologne, Herbert-Lewin-Str. 6, 50931 Köln, Germany
[b] University of Glasgow, School of Critical Studies, 4 University Gardens, Glasgow G12 8QQ, Scotland, United Kingdom

Received 5 July 2010; received in revised form 7 March 2013; accepted 11 March 2013
Available online

Abstract

Two perception experiments were conducted with subjects from Kiel, Düsseldorf and Vienna to investigate the role prosody plays (a) in resolving scope of negation ambiguities and (b) in judging the strength of phrasal breaks in German. The prosodic means tested were *pause*, *intonation contour* and *peak alignment*. Results reveal that the relevance of the cues varies depending on the task: for the (semantic) scope disambiguation task, *intonation contour* proves to be the most decisive factor, whereas presence of *pause* turns out to be most influential for the (metalinguistic) phrasing task. This result implies that the question of how German listeners resolve scope ambiguities cannot simply be attributed to the presence or absence of a phrasal break between a main and a subordinate clause. It rather seems to depend on a more general perception of 'cohesion' between the two clauses as indicated by prosodic means. Flat hat contours and late peak alignment patterns lead to a higher level of cohesion and an increase in wide scope interpretations, whereas pointed hats with early peak accents are typical of narrow scope readings. The results further reveal a significant difference between the varieties due to an increased number of narrow scope readings in Viennese listeners. Since Viennese German displays later peaks than Northern varieties, this outcome suggests that Viennese subjects interpret (late) peaks as earlier than listeners from Kiel and Düsseldorf.

Keywords: Scope of negation; Prosody; German varieties; Perception; Peak alignment; Phrasing

1. Introduction

Intonation is used in many languages to disambiguate syntactically and semantically ambiguous utterances (Jackendoff, 1972; Bolinger, 1989; Price et al., 1991; Avesani et al., 1995; Hirschberg and Avesani, 2000). One such type of ambiguity is a structure that allows for different interpretations as to the scope of negation, i.e. whether the negation applies to a constituent smaller or larger than a single clause,[1] giving rise to *narrow* and *wide scope* readings of a sentence.

The main motivation for the present paper was to do the first *perception* study on the prosodic parameters that influence the interpretation of scope of negation ambiguities in German. So far, the prosody of negation scope has only been investigated in production studies in a variety of languages. The relevant aspects of these studies for the present investigation will be presented in the following section (section 1.1). Furthermore, since German varieties are claimed to differ in their tonal alignment properties – again proposed in production studies only – we tested whether these differences had an impact on the perception of negation scope for listeners from three varieties of German (spoken in Kiel, Düsseldorf and Vienna). Section 1.2 gives an overview of alignment studies on German. The first two sections of the introduction lay the foundation for the research questions posed in the present study, elaborated in section 1.3.

* Corresponding author. Tel.: +49 221 470 4259; fax: +49 221 470 5938.
E-mail addresses: stefan.baumann@uni-koeln.de (S. Baumann), Tamara.Rathcke@glasgow.ac.uk (T. Rathcke).

[1] In the default case, a single clause is the domain of a negation quantifier in English (Crystal, 1992).

0024-3841/$ – see front matter
http://dx.doi.org/10.1016/j.lingua.2013.03.004

Section 2 reports on two perception experiments, consisting of a semantic task in which subjects had to judge whether a prosodically manipulated stimulus had a narrow or a wide scope reading (Experiment I) and a metalinguistic/prosodic task in which subjects had to judge the strength of a perceived intra-sentential phrase boundary (Experiment II). In section 3, we discuss the findings of our experiments and possible implications for the interface between prosody and meaning, before we sum up the results in a brief conclusion (section 4).

1.1. Prosody and the scope of negation

There has been some debate on the question of which prosodic cues contribute to the interpretation of negation scope in ambiguous sentences. An early (impressionistic) account for English has been presented by Jackendoff (1972:357f.), who discusses the difference between the two possible interpretations of the sentence *Max doesn't beat his wife because he LOVES her* (in which the main clause is presupposed and the subordinate clause is in focus). Here, meaning differences as to the scope of negation are claimed to be marked by different intonation contours which Jackendoff calls *A accent* and *B accent*[2]:

(1) Max does not beat his wife because he LOVES her.

'A accent', narrow scope

Paraphrase: 'Max does *not* beat his wife, the reason being mentioned'

(2) Max does not beat his wife because he LOVES her.

'B accent', wide scope

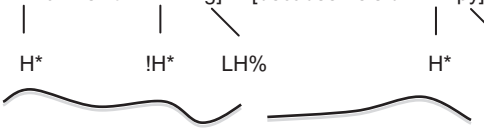
Paraphrase: 'Max *does beat* his wife, but not for the reason mentioned'

According to Jackendoff, a B accent (with a falling-rising contour) triggers the wide scope interpretation, i.e. the negation has scope over the whole sentence (paraphrase: 'Max *beats* his wife, but not for the mentioned reason'). In contrast, an A accent (with a falling contour) triggers the narrow scope interpretation, i.e. the negation has scope over the main clause only (paraphrase: 'Max does *not* beat his wife, the reason being mentioned'). In the framework of Autosegmental-Metrical Phonology (AM, e.g. Ladd, 1996/2008), these examples differ by the choice of utterance-final boundary tone, not the type of pitch accent.

In a more recent production study on ambiguously negated sentences in English and Italian, Hirschberg and Avesani (2000) discuss a variety of prosodic parameters which can be used for semantic disambiguation. An earlier and smaller-scale study by Avesani et al. (1995) additionally investigated scope ambiguities in Spanish. An English example from Hirschberg and Avesani (2000) is *William isn't drinking because he's unhappy* (for Italian, a target sentence with the same meaning was chosen). This utterance allows both for a wide scope reading according to which William drinks, but not because he is unhappy, and for a narrow scope reading in which William does not drink, the reason being his unhappiness (Hirschberg and Avesani, 2000:88). The main finding of the study is that English as well as Italian speakers disambiguate scope of negation not only by the choice of utterance-final boundary tone, but also, and in fact primarily, by intra-sentential phrasing (see also House, 2006:1549). The same strategy has been found for Spanish speakers by Avesani et al. (1995). In English, wide scope is usually indicated by a single phrase ending in a falling-rising movement or continuation rise, equivalent to Jackendoff's B accent, cf. examples in (2) and (3), whereas narrow scope is marked by an internal phrase boundary and a falling movement at the end of the whole utterance, equivalent to Jackendoff's A accent, cf. examples in (1) and (4).

(3) [WILliam isn't DRINking because he's unHAPpy]


wide scope

(4) [WILliam isn't DRINking] [because he's unHAPpy]


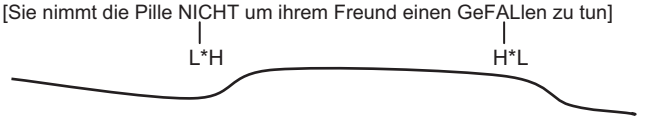
narrow scope

[2] Following Bolinger (1965), Jackendoff defines the term *pitch accent* as subsuming not only the accented syllable (as we will use the term for the rest of the paper) but also the following unaccented syllables up to the next accented syllable or the end of the utterance (see Jackendoff, 1972:258).

Italian speakers differed from the English speakers in the type of pitch accent they preferred (H+L* nuclear accents) and the strength of the boundary produced in the narrow scope versions. In the wide scope versions, they placed the nucleus on the verb of the main clause (*beve* 'drink') and deaccented the rest of the sentence. That is, Italian speakers use accentuation (or lack thereof) as one means of disambiguating the scope of negation. Very similar results for Spanish speakers were reported in Avesani et al. (1995) using a comparable read speech corpus.

Intra-sentential phrasing has not only been proposed to be relevant for scope disambiguation in English, Italian and Spanish but also for German. Féry's (1993) production data of ambiguous sentences such as *Sie nimmt die Pille nicht, um ihrem Freund einen Gefallen zu tun* ('She does not take the pill to please her boyfriend') indicate that the distinction between the two scope readings stems from phrasing differences accompanied by or even due to the choice of pitch accent type on the negation particle. According to Féry, the falling accent on *nicht* in (6) leads to the perception of a phrasal break,[3] which may be supported by temporal cues, in particular by a pause after the main clause. Additionally, the (local) accent types are integrated into different intonation contours or 'tunes' (in a holistic sense; see Dainora, 2006) marking either wide or narrow scope of negation. Consider the following examples from Féry (1993:141): in (5), a high plateau between the accent peaks marks wide scope,[4] whereas in (6), an F0 drop between the peaks marks the narrow scope reading.[5]

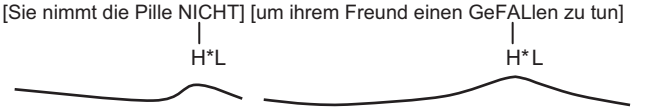
(5) [Sie nimmt die Pille NICHT um ihrem Freund einen GeFALLen zu tun]



wide scope

'It is not to please her boyfriend that she takes the pill' (She does not want a baby)

(6) [Sie nimmt die Pille NICHT] [um ihrem Freund einen GeFALLen zu tun]



narrow scope

'She does not take the pill, in order to please her boyfriend' (Her boyfriend wants a baby)

These two sentence-level tunes have been proposed in the IPO framework of intonation research ('t Hart and Collier, 1975; 't Hart et al., 1990; Adriaens, 1991; De Zitter, 1992). The contour in (5) has been commonly referred to as *flat hat* (or *hat pattern*), displaying a succession of a rise on the first accented syllable, followed by a plateau and a fall on the second accented syllable. In contrast, the contour in (6) shows accentual rises immediately followed by falls which have been called *pointed hats*. We adopt the IPO terminology here since it intuitively reflects the shape of the contours. AM-accounts of both contours have been discussed by Gussenhoven (1983), Wunderlich (1988), Selkirk (1995), or Welby (2003).

Since the present paper (including the experiments presented in section 2) is primarily concerned with German, we will base our study on the set of prosodic parameters discussed in Féry's (1993) work. Note, however, that she did not investigate the role of utterance-final boundary tones for German (as Hirschberg and Avesani did for English and Italian). Thus, for the sake of completeness, we conducted a small-scale perception experiment in addition to the present study, which revealed that the choice of utterance-final boundary tone has a strong influence on scope interpretation in German as well. In fact, 20 native listeners of Standard German interpreted the presence of a phrase-final rise as indicative of wide scope in about 80% of the cases (other prosodic parameters were varied as well), making it an important factor for scope interpretation (see results of the experiment in Appendix). For the rest of the paper, we will only deal with the factors governing differences in *intra-sentential* phrasing.

An important aspect of Féry's (1993) account on negation ambiguities in German is the fact that she treats both the (global) intonation contours *and* their (local) building blocks, namely the pitch accents the tunes are composed of, as

[3] According to GToBI (see Grice et al., 2005), Féry's H*L accent would translate into a combination of a H* pitch accent plus a L boundary tone (L- or L-%). Thus, the GToBI notation directly expresses the presence of a phrasal break.

[4] Féry's definition of this pattern as marking ''two phrases that are syntactically and semantically tied together'' (1993:151) expresses a strong *cohesion* between the phrases. The notion of 'cohesion' will be used in a broad sense in the course of this paper, expressing the (prosodically indicated) 'connectedness' or 'belonging together' of two phrases. Note that this definition deviates from the common usage following Halliday and Hasan (1976), in which cohesion is a property of text applying to a domain larger than a single sentence.

[5] The same prosodic means have also been claimed to be crucial for the differentiation between restrictive and appositive relative clauses in German (the former being marked by no pause and a high plateau between main and relative clause, the latter by a pause and falling intonation between the clauses; e.g. Seiler, 1960). Restrictive relative clauses narrow down the meaning of a specific element and are thus essential for its identification, whereas appositive relative clauses only provide additional (i.e. non-essential) information. This semantic difference entails different degrees of cohesion between main and relative clause (stronger cohesion in restrictive relative clauses), which is expressed – at least to some extent – by prosodic means (Schaffranietz, 1997).

essential for an utterance's semantic-pragmatic interpretation. If we consider different accent types as 'phonologised' differences in the *temporal alignment* of tonal targets with an accented syllable, we can hypothesise that peak alignment in general is an important cue in dissolving scope ambiguities. This seems to be plausible since alignment differences have been shown to play a crucial role in signalling information structure (e.g. Frota, 2002; D'Imperio, 2001a; Prieto et al., 2005; Baumann et al., 2007). However, what remains unclear is the question as to what extent and in which combinations variation in peak alignment can influence the interpretation of a sentence whose negation scope is ambiguous. In fact, if peak alignment (within a tune containing several accents) is a relevant cue for the correct interpretation of an ambiguously negated utterance in German, it is crucial to look at different varieties of this language as well, since German dialects have been shown to differ in the temporal alignment of accentual peaks. One of the major aims of this paper was to shed light on this issue by means of perception experiments. Before we turn to the empirical part (section 2), however, we will briefly present the basis for alignment studies in German and take a closer look at alignment differences in German varieties (section 1.2).

1.2. Peak alignment in German

The alignment differences that will be dealt with here are based on Kohler's (1991) distinction between 'early', 'medial' and 'late peaks'. In an early peak contour, the F0 peak is reached in the syllable immediately preceding the accented one, followed by a considerable drop in pitch onto the accented syllable. This contour can be analysed as an H+!H* or H+L* pitch accent in GToBI (Grice et al., 2005). The medial peak contour is characterised by an F0 peak around the centre of the accented syllable, accompanied by lower F0 values before and after this syllable. We assume that this pattern is equivalent to an H* pitch accent. Finally, in a late peak contour, a considerable part of the accented vowel is rather low and the peak is reached at the very end of the syllable. In GToBI terms, this contour can be transcribed as L+H* (or, if the peak is reached on the following syllable, as L*+H). The schematic contours of the sentence *Sie hat ja gelogen* ('She actually lied') in (7) are adapted from Kohler (1991:119; see also Grice and Baumann, 2002:290f.; Kohler, 1995:123)[6]:

(7)

	a.	b.	c.
Kohler (1991)	'early peak'	'medial peak'	'late peak'
GToBI	H+L*	H*	L+H* (or L*+H)

Kohler (1991) found a categorical change in perception between early and medial peaks, but not between medial and late peaks (testing Northern German subjects). That is, in autosegmental-metrical terms, he found evidence for a distinction between H+!H* or H+L* on the one hand and H* on the other, but not between H* and L+H*. These results support the pragmatic interpretation that early peaks mark given and medial peaks new information, whereas a further delay of the peak beyond the medial position is used to signal unexpectedness or emphasis, a function usually regarded as non-linguistic which does not necessarily have to be expressed phonologically.[7]

When talking about alignment differences and their potential phonological categorisation, it is important to distinguish between studies on prenuclear and nuclear accents. *Prenuclear* accents on topic initial material tend to have earlier peaks than on topic medial material in English (Wichmann et al., 2000). For German, contrastive topics have later peaks than non-contrastive ones (Braun, 2006, 2007; Braun and Ladd, 2003). These alignment differences are assumed to be gradient rather than categorical. However, there are a number of studies on *nuclear* accent peaks and their role in information structure. It has been shown that early peaks are used on given or semi-given material, whereas medial (and late) peaks are used on new material (Röhr and Baumann, 2011; Baumann, 2006; Baumann and Grice, 2006; Kohler, 1991, 1995 for German; Pierrehumbert and Hirschberg, 1990 for English). Furthermore, early peaks have been claimed to evoke a stronger impression of finality than medial or late peaks (Swerts et al., 1994; Wichmann et al., 2000). Thus, this distinction can be assumed to be phonological and categorical.

[6] The shaded areas indicate accented syllables.

[7] However, recent studies by Niebuhr (2003, 2007) show that the change from medial to late peak perception appears to be categorical if not only peak alignment but also other phonetic factors are taken into account (e.g. slope of the F0 rise and fall, peak height, durational and intensity relations of segmental constituents). Similarly, for English, both early vs. medial and medial vs. late distinctions have been found to be categorical. Pierrehumbert and Steele (1989) and Redi (2003) found a categorical distinction between L+H* and L*+H, and H+L* and H*, respectively, both using an imitation task. The nature of the task, however, precluded the possibility of assigning any functions to the categories. Similar findings have been presented for Italian (D'Imperio and House, 1997; D'Imperio, 2001b) and Russian (Makarova, 2007; Rathcke, 2006a,b).

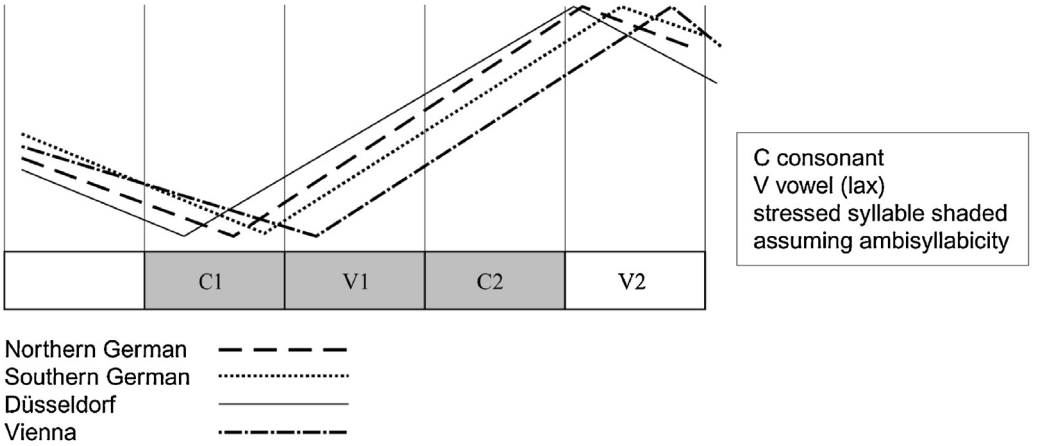


Fig. 1. Summary of alignment properties of prenuclear LH accents.
Illustration of Northern and Southern German varieties adapted from Atterer and Ladd (2004); Düsseldorf and Vienna added from Mücke et al. (2008).

1.2.1. *Varieties of German*

Previous studies on the variation in peak alignment between different dialects in German are neither consistent in the direction of alignment differences, nor do they conclusively answer the question how the temporal inter-dialectal variability should be described, namely in terms of a phonological pitch accent type or as gradual alignment differences. Nevertheless, all of these studies suggest that there *are* dialect-specific effects on peak alignment of some kind.

As early as in 1903, Eduard Sievers claimed that Low and High German have distinct intonation systems within the German language area. He further claimed that the intonation of one system is the mirror image of the other, i.e. where the pitch in one variety rises, it falls in the other one and vice versa. Although these claims are certainly too strong, there clearly are differences between Northern and Southern varieties. It has been reported that Southern German (including Austrian) varieties have later-timed accent peaks than Northern German varieties (e.g. Féry, 1993; Gibbon, 1998). This delay seems to be especially pronounced in the Vienna variety. A comparison with non-delaying varieties, such as the Düsseldorf variety spoken in the Mid-West of Germany and the Kiel variety spoken in the very North, i.e. in the Baltic Sea region of Germany, appears particularly fruitful. In fact, there is so far no consensus as to whether the variation in peak alignment across these varieties should rather be judged as phonetic or indeed phonological.

Evidence for different alignment properties of intonational peaks in the German language area has recently been provided by several acoustic and articulatory experiments. Ulbrich (2002) found in a cross-dialectal study of declarative utterances produced by speakers of Standard varieties of Germany, Austria and Switzerland that German speakers make use of falling contours with a relatively early aligned peak on the nuclear syllable, whereas Austrian (Viennese) speakers produce late-peak falls. Results of production experiments with speakers from Northern and Southern Germany (Atterer and Ladd, 2004) and from Düsseldorf and Vienna (Mücke et al., 2008) also confirmed the claim that accent peaks occur later in Southern varieties than in Northern varieties. Both studies investigated bitonal rising prenuclear pitch accents in non-contrastive contexts.

Fig. 1, which combines the results of the two studies, illustrates the similarity of the alignment pattern between the Düsseldorf variety and Atterer and Ladd's 'Northern German' variety. In both varieties, L is aligned with the onset consonant of the accented syllable (C1), and H is aligned with the beginning of the vowel of the following syllable (V2).[8] Atterer and Ladd's 'Southern German' and the variety spoken in Vienna both display a later alignment of L and H than the Düsseldorf or Northern German varieties. Although both H and L are later in the Vienna variety, the alignment of L is different in a crucial way: whereas in Southern German L is late – but like the other varieties, still in the onset consonant – in Vienna, it shifts into the vowel (V1).

However, Braun (2007) showed that prenuclear rises are aligned later in Southern German (as spoken in Munich) than in Northern German (as spoken in Münster) only in sentences with a high nuclear accent (H*), produced in non-contrastive contexts. In fact, Braun claims that the alignment differences between Northern and Southern German speakers are conditioned by both pragmatic (marking of contrast) and intonational context (type of nuclear accent).

[8] Fig. 1 is stylized in that durational differences between the segments in the dialects were normalized for.

Table 1
Overview of accent properties found in acoustic alignment studies (Grice and Harrington, 2003; Mücke et al., 2006, 2008, 2009) on the varieties spoken in Kiel, Düsseldorf and Vienna.

	Kiel	Düsseldorf	Vienna	
Prenuclear, contrastive	–	Late peak	Late peak	Same category, later peak in Vienna
Prenuclear, non-contrastive	–	Late peak	Late peak	Same category, later peak in Vienna
Nuclear, contrastive	–	Late peak	Late peak	Same category, later peak in Vienna
Nuclear, non-contrastive	Medial peak	–	Late peak	Different category, late peak in Vienna

In a set of acoustic and articulatory experiments, Mücke and colleagues (Mücke et al., 2006, 2009), investigated contrastive L+H accents both in prenuclear and nuclear positions with speakers from Düsseldorf and Vienna. The acoustic analyses confirm later peak alignment in Vienna both in prenuclear (as opposed to the findings in Braun, 2007) and nuclear position. However, the alignment difference across varieties was only small. In fact, the H peak occurred within the same segment in both varieties – the postaccented vowel in prenuclear accents and the intervocalic consonant in nuclear accents.

This slight F0 delay found in the Vienna variety as opposed to the Düsseldorf variety deviates from findings of a preliminary (acoustic) comparison of Viennese and Kiel speakers by Grice and Harrington (2003). The authors found that the F0 peak in nuclear accents of declarative utterances was considerably later and higher in the Viennese data examined in comparison to the Kiel data. These results led them to suggest a categorical distinction between the nuclear accent types used in the two varieties: H* (medial peak) in Kiel versus L*+H (late peak) in Vienna.

Since *perceptual* evidence for alignment differences between German varieties is lacking, we set up experiments (see section 2) in order to test whether and to what extent the reported alignment differences have an influence on the (semantic as well as phonetic) interpretation of specific intonational tunes by listeners from Kiel, Düsseldorf and Vienna. Apart from the fact that these varieties have already been subject to studies on peak alignment providing us with the basis for our experimental hypotheses (Table 1 gives an overview of the findings presented in this section), our decision to investigate data from Kiel and Vienna is also based on the circumstance that the two varieties of German represent extreme poles of the North-South divide mentioned above. The third variety included, that spoken in Düsseldorf, is claimed to be closer to the standard variety of Northern German. This variety lacks the peculiarities of the coastal dialect region and is located in an area just north of the Benrath isogloss, which divides Low German (spoken in the North) from High German (spoken in the South).

1.3. Research questions

So far, we have presented an overview of a number of prosodic parameters that have been reported in production studies on various languages to disambiguate the scope of a negation particle. These were *utterance-internal* and *utterance-final boundary tone* (especially in English), *pitch accent type* or *peak alignment* (Italian, Spanish, German), *global intonation contour* (German), *accent placement* (Italian, Spanish) and *pause* (German). For all languages mentioned it has been claimed that these prosodic cues (except for utterance-final boundary tone) also have an impact on *intra-sentential phrasing*. This suggests that phrasing is the actual cue to resolving scope of negation ambiguities, and that the single parameters only add to the impression of a phrasal break or, respectively, lack of a break. In two experiments, we investigate the relation between both aspects, i.e. the role of several prosodic parameters and their effect on the perception of a phrasal break.

Since our experiments are on German, we concentrate on the parameters proposed for German by Féry (1993), i.e. *intonation contour*, *peak alignment* (or pitch accent type) and *pause*, although this is not to claim that some alternative parameters found for other languages are necessarily irrelevant for German.

The two experiments deal with three varieties of German, spoken in Kiel, Düsseldorf and Vienna. As mentioned above, production studies revealed (slight) alignment differences between these varieties but it remained unclear whether listeners *perceive* these distinctions differently. We are therefore seeking to explore whether different dialectal alignment properties (among the other prosodic cues proposed by Féry) are reflected in different semantic interpretations (scope of negation; Experiment I). In order to find out whether the prosodic parameters manipulated in Experiment I actually reflect degrees of intra-sentential *phrasal breaks*, we conduct another, purely prosodic (or metalinguistic) test (Experiment II).

In sum, the perception experiments described in section 2 are meant to answer the following research questions:

- Which of the prosodic parameters *intonation contour* (flat hat pattern vs. pointed hats), *peak alignment* (early, medial and late peaks) and *pause* (presence or absence of a pause) influence the semantic interpretation as to the scope of negation (Experiment I)?

- To what extent do the three parameters *intonation contour*, *peak alignment* and *pause* influence the perception of an intra-sentential phrasal break (Experiment II), and does the metalinguistic/prosodic judgment (i.e. strength of a perceived boundary) reflect the semantic judgment (i.e. wide or narrow scope of negation)?
- Do listeners of the three German varieties differ in their perception and functional interpretation of the prosodic parameters?
- Do the perceived alignment differences lead to different categorical distinctions (i.e. accent types) in the three varieties? Are (global) tunes or specific accent types (as the local building blocks of tunes) central for conveying meaning? To what extent do they interact?

2. Experimental data

In Experiment I, the subjects' task was semantic in nature, since they were asked to decide in a forced-choice test whether they interpreted the realisations of the German test sentences either as 'yes' or 'no' readings depending on the perceived scope of negation (see section 2.2). The experiment was further designed to test whether there are dialectal differences in the prosodic disambiguation of scope of negation sentences. We conducted the semantic task and an additional silent reading task with subjects from Kiel, Düsseldorf and Vienna. The stimuli were kept constant in terms of accent position (accents on negation particle and complement) and sentence-final intonation contour (low boundary tone), while the parameters *pause*, *intonation contour* and *peak alignment*[9] were systematically varied.

Experiment II, in which we used the same set of manipulated stimuli as in the first experiment, addressed a research question that was metalinguistic in nature as it targeted at an explicit judgment of prosodic structure. This study investigated which parameters influence the perception of a break (i.e. the degree of 'disconnectedness') between a main clause and a subordinate clause – and to what extent.

In the next section (section 2.1), we present the hypotheses to be tested in the experiments, before describing the selection and manipulation of the stimuli used (section 2.2). The last two sections deal with the methods employed and the results obtained in Experiment I (section 2.3) and Experiment II (section 2.4).

2.1. Hypotheses

Based on the literature review presented in the introduction, our general hypothesis is that a listener's impression of a strong phrasal break between the main and the subordinate clause leads to a narrow scope reading whereas the lack of a break triggers a wide scope reading. The prosodic parameters we consider to have an effect on both types of perception (the semantic task of Experiment I and the metalinguistic task of Experiment II) are the presence or absence of a pause, the shape of the intonation contour, and the alignment of accentual peaks. Moreover, we expect to find differences in the judgments of listeners from the three varieties, stemming from the alignment differences demonstrated in production studies of the respective varieties.

2.1.1. Hypothesis 1: Pause

A silent pause after the main clause (accompanied by pre-final segmental lengthening) leads to the perception of a phrasal break (e.g. Peters, 2007), which, in turn, triggers a narrow scope reading. In contrast, the lack of a pause leads to a wide scope reading (Féry, 1993).

2.1.2. Hypothesis 2: Intonation contour

A sequence of two pointed hats is expected to induce the perception of a phrasal break, due to the L tone between the accentual peaks. By contrast, a flat hat pattern is assumed to be perceived as a single phrase (following Welby, 2003). Thus, we expect a sequence of two pointed hats to trigger the narrow scope interpretation and the flat hat pattern to trigger the wide scope interpretation (cf. Féry, 1993).

2.1.3. Hypothesis 3: Peak alignment

We hypothesise early peak accents in the *main* clause (i.e. on the negation particle) to trigger the perception of a break between the two clauses due to an impression of givenness and, more importantly, finality of early peaks as opposed to medial and late peaks (Kohler, 1991; Swerts et al., 1994; Prieto et al., 1995; Wichmann et al., 2000). In contrast, late peak accents in the main clause are not expected to evoke the perception of a phrasal break since they often occur in prenuclear position (Wichmann et al., 2000; Atterer and Ladd, 2004). Consequently, we expect that early peaks in the main clause

[9] We deliberately use the term 'peak alignment' instead of 'accent type' in order to avoid determining the tonal categories which may vary in the dialects investigated. The question of accent type will be addressed in section 3.

support narrow scope interpretations, while late peaks trigger wide scope interpretations. The less marked medial peak accents are assumed to evoke either narrow or wide scope interpretations (they serve as control stimuli in the study).

The accent types used in the *subordinate* clause are assumed to be less crucial for scope interpretations, since they have no impact on a potential intra-sentential break. According to the production data by Braun (2007),[10] however, early peaks in the subordinate clause are expected to evoke more wide scope interpretations if they are part of a flat hat pattern, presumably adding to the perception of the contour as a cohesive whole. In contrast, medial and late peak accents in the subordinate clause are assumed to trigger the perception of a narrow scope of negation, especially if they are part of a sequence of two pointed hats. This is because these accent types, which are claimed to encode new information (Kohler, 1991), may be interpreted to mark a self-contained bit of additional information in this intonational context, rather than suggesting a lack of cohesion between main and subordinate clause. Thus, apart from a main effect of the factor *peak alignment* we expect an interaction between *alignment* and *contour* (for Experiment I).

2.1.4. Hypothesis 4: Dialectal influence

The three parameters *pause*, *intonation contour* and *peak alignment* are used as perceptual cues for phrasing and for resolving scope ambiguities in all three varieties under investigation. However, we expect systematic variation in the interpretation of peak alignment differences, given the fact that accent peaks in the Vienna variety usually occur later than in the varieties spoken in Düsseldorf or Kiel (Mücke et al., 2006, 2009). Thus, Viennese listeners should generally perceive accent peaks as earlier than listeners from Kiel and Düsseldorf, which, in turn, may result in an increase in narrow scope interpretations for Viennese subjects. That is, we expect an interaction between the factors *alignment* and *variety*.

2.2. Stimuli

The experiments were based on the test sentence *Sie hat den Roman NICHT gelesen, weil er auf ENGlisch ist* ('She did not read the novel because it is in English'), which is ambiguous as to the scope of its negation: a 'yes' reading ('She read the novel, but for another reason'; wide scope) contrasts with a 'no' reading ('She did not read the novel, the reason being that it is in English'; narrow scope). The test sentence was spoken by a trained male speaker of Standard German (Northern variety) in two versions, i.e. with a flat hat contour and with pointed hats. Both versions were produced with a low utterance-final boundary tone. The pointed hats version served as the basis for stimuli resynthesis since flat hat realisations often display segmental reduction between the two accent positions (as found by Peters et al., 2005). Furthermore, we recorded a filler sentence with a syntactic structure similar to the test sentence, and resynthesized it along the same lines as the test sentence.

We decided to use one speaker from a single dialect in order to be able to use the same materials for all three varieties. We assumed that segmental differences between the varieties could be neglected for two reasons: (1) the listeners' tasks were non-segmental in nature, i.e. they were asked to judge the speech melody of the test sentences (as indicated in the questionnaire on which the subjects marked their judgments), and (2) we used an off-line experimental procedure and gave all participants sufficient time for decision (cross-dialectal phonetic differences have been shown to slow down reaction times but otherwise do not affect judgments; see e.g. Floccia et al., 2009).

In line with the hypotheses, three prosodic parameters of the base sentence were manipulated using *Praat* software (Boersma and Weenink, 1996). An overview of the experimental design is given in Table 2. Fig. 2 visualises the acoustic manipulations of the stimuli. The actual Hz values of the peaks and valleys in the F0 trajectory (including a slight declination of top and base lines) are taken from the original productions of the test utterance.

The manipulation design described above led to 20 test stimuli (2 pause conditions $\times$ 2 intonation contours $\times$ 5 peak alignment combinations). There were two reasons for not including all possible peak alignment conditions (i.e. we left out the medial-early and medial-late combinations in both directions). First, we aimed at reducing the number of test stimuli to a minimum for each subject since we planned to present an equal number of stimuli and fillers in three repetitions (which would have led to 216 items plus practice stimuli instead of the 120 items plus practice stimuli we used in Experiments I and II). Second, medial peaks were included as controls so that the test stimuli combining late and early peaks would be evaluated against the default m/m-condition, giving us some relevant baseline information.

It is important to note that the nature of the experimental design (aiming at a comparison between flat and pointed hats in a systematic way) led to resynthesis artefacts in some specific cases: e.g. if a flat hat starts off with an early peak accent, the acoustic manifestation of the L is missing, since a flat hat requires a (more or less) linear interpolation between two peaks. In cases like this, the intended accent shape is only indicated by a significantly earlier H target (in comparison to a

[10] Braun (2007) found an increase in early peak accents (in nuclear position) in contrastive utterances, which can be interpreted as equivalent to our wide scope of negation readings (see Discussion). However, the sentences used by Braun were much shorter, i.e. they did not include subordinate clauses.

Table 2
Overview of manipulated parameters for the different test factors and conditions.

Factor	Condition	Manipulation
Pause	No pause	No silent pause between main and subordinate clause
	Pause	Silent pause of 300 ms between main and subordinate clause accompanied by 20 ms lengthening of final syllable rhyme (*-lesen*)
Intonation contour	Flat hat	Linear interpolation between accent peaks
	Pointed hats	Low F0 target between accent peaks (aligned with end of main clause)
Peak alignment	Early-early (e/e)	Succession of two accents with F0 maxima placed in preaccentual syllable (*-man, auf*) and low F0 target in accented syllable (*nicht, Eng-*)
	Early-late (e/l)	Accent with F0 maximum placed in preaccentual syllable (*-man*) and low F0 target in accented syllable (*nicht*) followed by accent with F0 maximum reached in postaccentual syllable (*-lisch*)
	Medial-medial (m/m)	Succession of two accent peaks with F0 maxima within the accented vowels (on *nicht* and *Eng-*)
	Late-early (l/e)	Accent with F0 minimum on accented syllable (*nicht*) and F0 maximum on postaccentual syllable (*ge-*) followed by accent with peak on preaccentual syllable (*auf*) and low target on accented syllable (*Eng-*)
	Late-late (l/l)	Succession of two accents with F0 maxima on postaccentual syllables (*ge-, -lisch*)

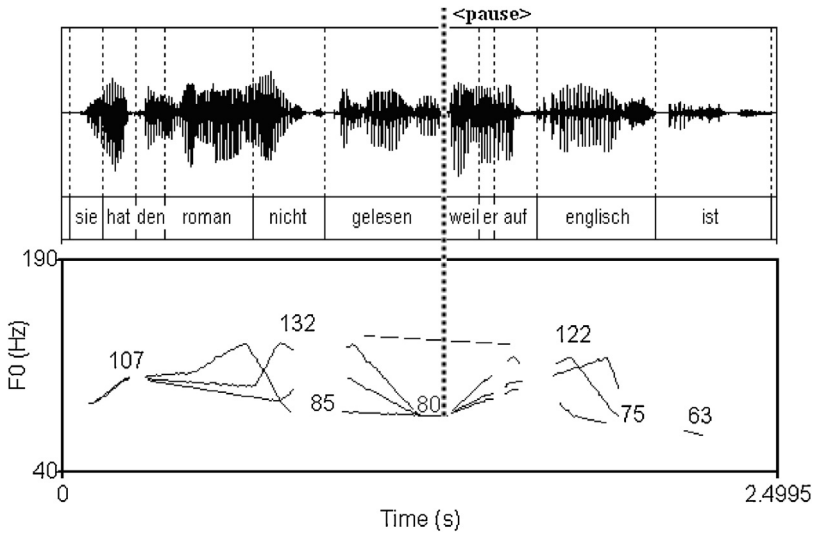


Fig. 2. Prosodic manipulations of the test sentence: actual Hz values are given for each relevant point of the trajectory; solid lines show F0 contours corresponding to pointed hats, the dashed line indicates linear interpolations between accentual peaks for flat hat contours; the dotted vertical line points at the place of pause insertion.

medial peak). These artefacts are another reason for our decision to refer to differences in *peak alignment* rather than to different *accent types*.

As for the *pause* condition, the phrase-final syllable rhyme was lengthened for 20 ms in order to achieve the final lengthening effect (Turk, 1999; Wightman et al., 1992) and thus to create more natural-sounding stimuli. However, the insertion of a pause in a flat hat contour might be regarded as counter-intuitive and thus unnatural. Nevertheless, Wunderlich (1988:12f.) and also Féry (1993) claim for German that a hat pattern (in Wunderlich's terms *Brückenakzent*) may well be interrupted by a pause. The general perceptual acceptability of the stimuli was ensured by informal judgments collected during stimuli preparation at Kiel University.

2.3. Perception Experiment I: scope ambiguities

The first experiment investigates (a) which prosodic cues influence listeners' interpretation as to the scope of negation in ambiguous sentences and (b) whether there are dialectal differences in these interpretations.

2.3.1. Subjects and method

Three groups of listeners with different dialectal backgrounds participated in the experiment: 17 participants from both Kiel (11 m, 6 f) and Düsseldorf (7 m, 10 f) and 18 participants from Vienna (14 m, 4 f). All of them were between 20 and 30 years old (mean 24.3) and did not report any hearing or speaking disorders. All listeners were paid for their participation.

The stimuli described above were presented over loudspeakers in a sound treated room and to small groups of listeners simultaneously. The 20 test stimuli and 20 fillers were repeated three times each and presented in randomised order. The resulting 12 blocks of 10 stimuli each were preceded by a training block and followed by a finishing block – those were not evaluated. Thus, the total number of experimental stimuli amounted to 140 (20 stimuli + 20 fillers × three repetitions plus 20 dummies in two additional blocks). One test session lasted approximately 30 min.

The data were obtained in a forced-choice manner, i.e. the listeners' task was to answer the question 'Did she read the novel?' for each stimulus by ticking a box on a paper questionnaire saying either 'yes' or 'no'. A 'yes'-response indicates a wide scope reading ('she read the novel, but for another reason'), whereas a 'no'-response ('she did not read the novel') indicates a narrow scope interpretation. After the stimulus presentation participants had three seconds for their decisions. This procedure seemed to be very clear and easy to follow for all participants as we did not encounter any missing responses.

An additional test was run with a different group of participants (15 per variety).[11] The task was to silently read several semantico-syntactically ambiguous sentences. One of them was our test sentence *Sie hat den Roman nicht gelesen, weil er auf Englisch ist* ('She did not read the novel because it is in English'). Subjects had to mark on a sheet of paper which of two given interpretations (here: A. She read the novel, but for another reason, B. She did not read the novel; the order of the answers was counterbalanced) was the most suitable description of the sentence's meaning. This task was added in order to find out whether there was a bias towards one of the two possible scope interpretations and whether there were differences between the varieties tested.

As to inferential statistics, all data were analysed using `lme4` and `languageR` packages in R (Bates and Maechler, 2009; Baayen, 2009). In Experiment I (and in the boundary tone experiment; see section 1.1 and Appendix), we used mixed logit models with a binomial error function and *subjects* as a random effect (see Jaeger, 2008) for the analysis of categorical dependent variables.

Generally, we validated the significance of fixed effects through comparing the test model (with fixed effects) to a reduced model (without fixed effects) in likelihood ratio tests (LRTs). Fixed effects that contributed to a better model fitting were included in the final model.

2.3.2. Results

Fig. 3 illustrates the mean values of 'yes'-responses (indicating the perception of the wide scope reading) for each prosodic parameter across all varieties.

Considering the main tendencies that can be derived from Fig. 3, *pause* only had a weak effect on the perception of the negation scope in all three varieties. Nevertheless, it became obvious that lack of pause supports wide scope readings. In contrast, the effect of *intonation contour* was very prominent, especially in the Kiel and Düsseldorf varieties. Flat hat contours increased the proportion of 'yes'-responses, i.e. wide scope interpretations. *Peak alignment* proved to be perceptually relevant as well, with early peaks inducing more narrow scope responses and later peaks more wide scope readings.

In order to test the effect of the manipulated parameters on perception by inferential statistics, we constructed a model with binomial error function using the scope judgments (wide vs. narrow) as a categorical dependent variable. We started out with a model including *subjects* as a random effect and included the following two-way interactions as fixed effects: the factor *variety* with three levels (Kiel, Düsseldorf, Vienna) interacting with the factor *intonation contour* with two levels (flat hat, pointed hats), *variety* interacting with the two-level factor *pause* (present or absent between two clauses), *variety* interacting with the factor *peak alignment* with five levels (e/e, e/l, m/m, l/e, l/l), *pause* interacting with *peak alignment*, *pause* interacting with *intonation contour*, and *peak alignment* interacting with *intonation contour*. We applied a backward fitting procedure starting with the general model above and reducing it stepwise to find the best fit via likelihood ratio tests (LRTs). LRTs revealed that only three interactions significantly improved the data fit: *variety* × *intonation contour* (LRT: $\chi^2(3) = 409.69$, $p < 0.001$), *variety* × *peak alignment* (LRT: $\chi^2(8) = 48.62$, $p < 0.001$), and *pause* × *peak alignment* (LRT: $\chi^2(5) = 28.48$, $p < 0.001$). No other interaction entered the model. Moreover, Wald's *z*-test indicated a significant main effect of *intonation contour* ($p < 0.001$), *peak alignment* ($p < 0.001$) and *pause* ($p < 0.01$). The final model (with fixed effects) fitted the data significantly better than a null model (without fixed effects) (LRT: $\chi^2(22) = 635.98$, $p < 0.001$).

The significant interaction between *variety* and *intonation contour* indicates that the Viennese subjects did not differentiate as strongly between a flat hat and pointed hats in their decision as to the scope of negation than subjects from

[11] This set of subjects also took part in Experiment II (see section 2.4).

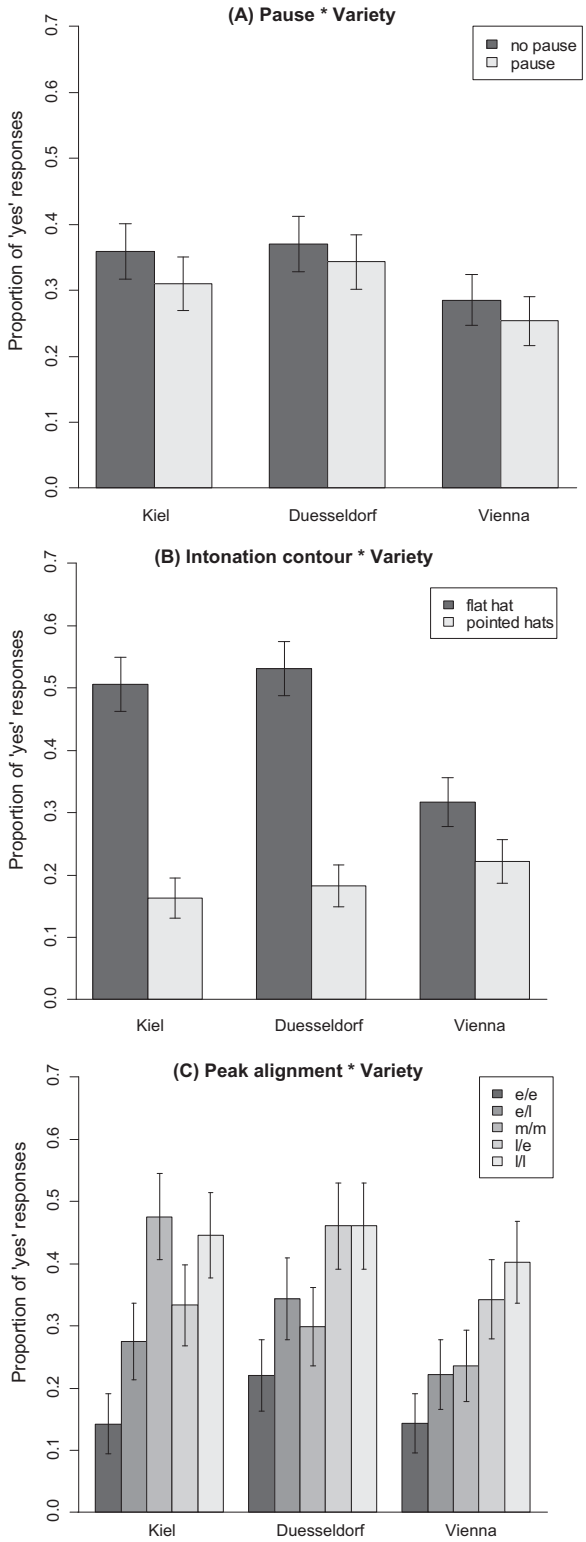


Fig. 3. Proportion means of 'yes'-responses for the factors *pause* (A), *intonation contour* (B) and *peak alignment* (C) across the three German varieties investigated (cf. Table 2 for explanation and coding of the conditions per factor). Error bars represent 95% confidence intervals.

the other varieties. That is, for both Kiel and Düsseldorf listeners, scope interpretation mainly depended on the shape of the *intonation contour* (cf. Fig. 3B). *Variety* also interacts with *peak alignment*. Broadly speaking, Kiel listeners judged early peaks as different from medial or late peaks, and Düsseldorf listeners judged early and medial peaks different from late peaks. In fact, the tendency of early peaks to trigger narrow scope was less obvious than in Kiel. On the other hand, accent peaks with late alignment more consistently led to wide scope interpretations in the Düsseldorf variety. The picture is less clear with subjects from Vienna (but see below and the Discussion for a possible explanation). Nevertheless, the relative influence of late peak alignment on the perception of wide scope proved to be strong as well (cf. Fig. 3C). Finally, the interaction between *pause* and *peak alignment* indicates that the lack of a pause after the main clause more often led to a wide scope interpretation if it was accompanied by a late peak accent on the negation particle.

As a general result, Viennese subjects (whose variety utilises the latest peaks among the three varieties under investigation) interpreted the stimuli less often as encoding wide scope (average proportion of 'yes'-responses is 27%) than subjects from Kiel (33%) and Düsseldorf (37%). These numbers indicate that there was a general bias towards the narrow scope reading, since the average numbers for 'yes'-responses were below the arithmetic mean of 50% in all three varieties. This bias was clearly confirmed by the additional *silent reading* experiment mentioned in section 2.3.1 above: 93% of the Viennese subjects (i.e. 14 out of 15) and even all subjects from Kiel and Düsseldorf interpreted the written test sentence as having narrow scope. Obviously, the syntactic structure displaying a main clause and a subordinate clause which are separated by a comma, is more likely to trigger an interpretation of two separate pieces of information (see also Kentner, 2010; Hemforth and Koniecny, 2004). In cases like this, prosodic means have to be very salient in order to override the syntactically induced interpretation bias.

To sum up the main results of Experiment I, we found strong effects of *intonation contour* as well as *peak alignment* as factors used by our subjects to resolve scope of negation ambiguities, both showing interactions with *variety*. Generally, the effects were weaker for Viennese listeners. The factor *pause* turned out to be less important for the disambiguation task.

2.4. Perception Experiment II: phrasing

According to our basic hypothesis, the perception of a phrasal break between the main and the subordinate clause crucially determines the interpretation of a sentence as to the scope of its negation. Thus, we want to find out whether the prosodic cues which proved to be relevant for the semantic decision as to the scope of the negation also have an impact on the (purely prosodic) perception of phrasing, and to which extent. For this reason, we tested the same stimuli from Experiment I (see section 2.2) in a phrasing experiment.

2.4.1. Subjects and method

Fifteen subjects per variety (Kiel: 6 m, 9 f; Düsseldorf: 5 m, 10 f; Vienna: 5 m, 10 f), aged between 18 and 57 (mean 29.8), were tested. They did not report any hearing or speaking disorders and were paid for participation. None of the subjects participated in Experiment I. The listeners were asked how strongly disconnected they felt the two clauses of the test sentence to be.[12] The answers had to be marked on a paper questionnaire providing a nine-point scale for each utterance. Position 0 of the scale was labelled ''not disconnected at all'' and position 8 ''very strongly disconnected''. The subjects were encouraged to use the whole range of the scale during the experiment.

As in Experiment I, the stimuli were presented over loudspeakers in a sound treated room and to small groups of subjects simultaneously. The same 20 test stimuli and 20 fillers from Experiment I were presented in randomised order. The utterances were interspersed by pauses of three seconds in which the subjects had to make their decisions. The test began with a training block of 10 dummy stimuli and ended with a dummy block of the same length, resulting in a total of 60 stimuli. Each stimulus was presented only once. One complete session lasted about 15 min.

The statistical method was the same as in Experiment I, with the exception that the dependent variable in Experiment II is not categorical. Thus, we analyzed the *z*-transformed continuous dependent variable using linear mixed-effects models with *subjects* as a random effect (Baayen et al., 2008).

2.4.2. Results

First of all, the raw scale values were *z*-score transformed to standardise across individual differences in the usage of the categorical scale. The results below are based on the standardised scores. Fig. 4 shows the listeners' judgments as to the strength of the perceived phrasal breaks for each prosodic parameter across all varieties.

[12] The actual question in German was *Wie stark sind die beiden Teilsätze voneinander getrennt?* ('How strongly are both clauses disconnected?').

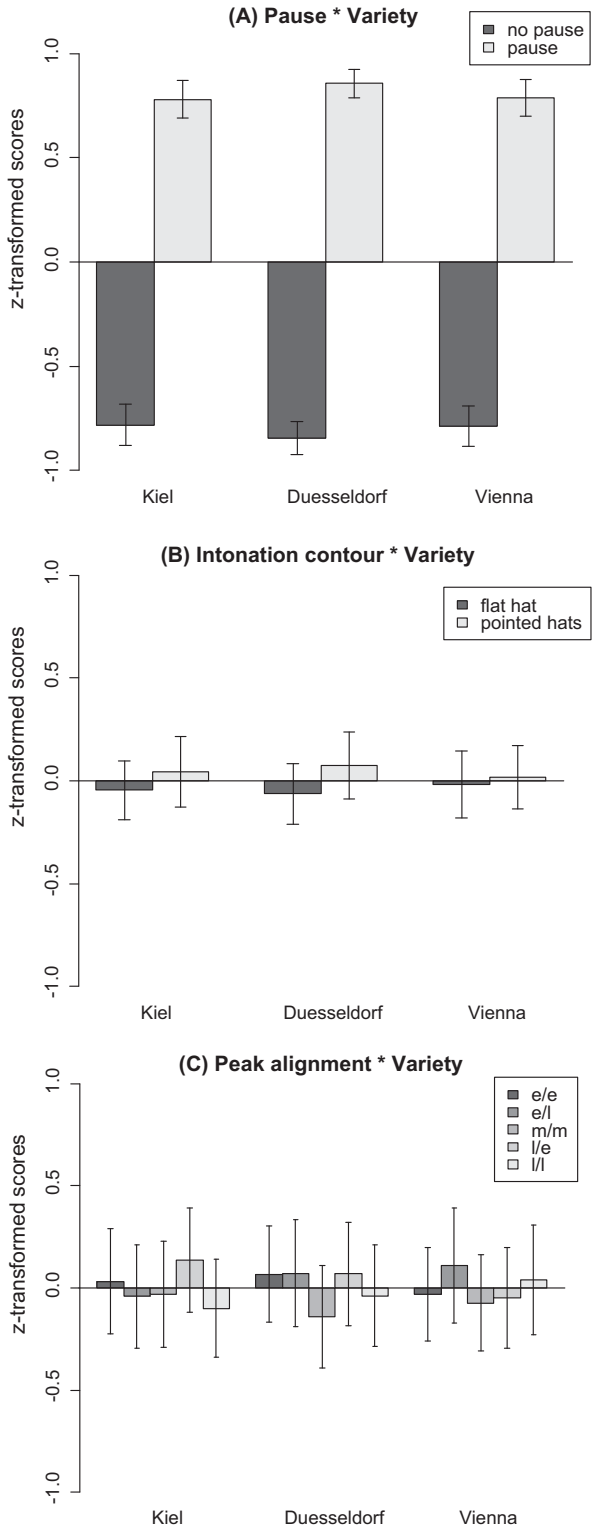


Fig. 4. Perceived strength of phrasal break (z-transformed scores) for the factors *pause* (A), *intonation contour* (B) and *peak alignment* (C) across the three German varieties investigated (cf. Table 2 for explanation and coding of the conditions per factor). Error bars represent 95% confidence intervals.

For the statistical analysis we used the *z*-transformed perceived phrasal break judgments as a (continuous) dependent variable. First, we tested a model using *subjects* as a random effect and included the following two-way interactions as fixed effects: the categorical factor *variety* with three levels (Kiel, Düsseldorf, Vienna) interacting with the categorical factor *intonation contour* with two levels (flat hat, pointed hats), *variety* interacting with the two-level categorical factor *pause* (with pause, without pause), *variety* interacting with the categorical factor *peak alignment* with five levels (e/e, e/l, m/m, l/e, l/l), *pause* interacting with *peak alignment*, *pause* interacting with *intonation contour*, and *peak alignment* interacting with *intonation contour*. As in Experiment I, we applied a stepwise exclusion of fixed effects which did not fit the data significantly better via likelihood ratio tests (LRTs). LRTs revealed that only the two-way interaction *pause* × *intonation contour* (LRT: $\chi^2(3) = 1066.1$, $p < 0.001$) could be kept as a significant predictor. No other interaction entered the model. Moreover, it did neither include *peak alignment* nor *variety* as main effects since they did not contribute to a significant improvement of the model fit (LRT *alignment*: $\chi^2(4) = 8.355$, $p = 0.08$; LRT *variety*: $\chi^2(6) = 8.382$, $p = 0.21$). However, based on Baayen's *p*-value generating function, the final model revealed significant main effects of *pause* ($p < 0.001$) and *intonation contour* ($p < 0.001$).

Apparently, variation in *peak alignment* was irrelevant for this task, and the factor *intonation contour* only played a minor role. In contrast, the presence or absence of a *pause* was the main cue for perceived phrasing in each variety. The factor *pause* showed the strongest effect of the analysis. Actually, the presence of a pause (accompanied by phrase-final lengthening) induced a particularly clear perception change in listeners from Düsseldorf (but also from the other varieties, see Fig. 4A), shifting the judgment from 'rather not disconnected' to 'quite strongly disconnected'. *Intonation contour* also affected the perception of a phrasal break significantly – with pointed hats leading to stronger breaks in all three varieties – although the magnitude of the perception change induced by this variable was comparably low (see Fig. 4B). Nevertheless, *intonation contour* interacted significantly with *pause* showing that pointed hats systematically improved the perceptual impact of a pause. In contrast, no significant effect of *peak alignment* (neither as a main effect nor in an interaction) could be observed.

Briefly summing up the main results of Experiment II, we can state that the perception of a phrase break was mainly determined by the presence or absence of a *pause*, showing the same strong effect in all three varieties. *Intonation contour* also had a significant but weak effect.

Finally, in order to find out whether there is a correlation between the results of both experiments, and thus whether there is an interdependence between scope of negation and phrasing, we calculated means for the dependent variables in Experiment I (proportions of 'yes'-responses) and Experiment II (*z*-transformed scores) for each prosodic factor combination ($2 \times 2 \times 5$). A Spearman rank correlation test did not reveal a significant correlation ($r_s(18) = -0.24$, $p = 0.31$).

3. Discussion

Our general hypothesis put forward in section 2.1 proved to be too strong in that we did not find a significant correlation between the perceptual cues used for scope interpretation and the ones used for the interpretation of phrasing. Rather, the relevance of a parameter was found to vary depending on the task, which was semantic in Experiment I but metalinguistic in Experiment II. Nevertheless, the parameters investigated here seem to serve a more general function which can be described in terms of *cohesion* or *connectedness*: We found that those prosodic parameters which enhanced the degree of cohesion between the main and the subordinate clause (especially a flat hat and lack of a pause) often triggered more wide scope interpretations and supported the perception of no phrasal break. However, the relation between scope interpretation and phrasing is less direct than expected. It would thus be wrong to say that the lack of a prosodic break generally triggers a wide scope reading or, conversely, that a phrasal break always triggers a narrow scope reading.

Such a direct relation has not been explicitly claimed by the production studies on several languages which we discussed in section 1.1 and on which we built our research questions. In fact, many of the parameters proposed to both trigger prosodic phrasing and play a role in scope disambiguation could be confirmed to be relevant for the perception of scope ambiguities in German, in particular global intonation contour (Féry, 1993) and peak alignment (Avesani et al., 1995; Hirschberg and Avesani, 2000; Féry, 1993). However, we have to acknowledge that the notion of 'phrasing' can be used in a broader sense, under which our notion of 'cohesion' could be subsumed, or in a stricter sense denoting the actual perception of an intonation phrase boundary. For the purpose of the present study, we used the term 'phrasing' in the stricter sense.

In the scope experiment (Experiment I), the factor *intonation contour* proved to be the most powerful one although *peak alignment* and (to a smaller extent) *pause* also had a significant influence on the perceived scope of negation. In the phrasing experiment (Experiment II), however, the parameter *pause* turned out to be the most influential cue for the perception of a phrasal break, supported by *intonation contour*. For this task, differences in *peak alignment* did not play a role. In terms of the dialectal influence, we only found perceptual differences in the judgments of the global intonation contour as well as the alignment patterns in Experiment I, but no effect of *variety* in Experiment II.

For a more detailed discussion of the three parameters investigated here, we will consider each hypothesis separately. The first three hypotheses are based on the results of all participants, while dialectal differences are discussed with respect to hypothesis 4.

Hypothesis 1: Pause. The hypothesis could generally be confirmed: A silent pause after the main clause (accompanied by pre-pausal segmental lengthening) proved to be the strongest indicator for a phrasal break (Experiment II, Fig. 4A). Furthermore, the presence of a pause significantly increased the proportion of narrow scope judgments across all subjects in Experiment I, suggesting a link between scope interpretation and phrasing. However, the actual effect of pausing on scope interpretation was only weak (Fig. 3A). One reason might be the predominance of the factor *intonation contour*, which reduces the effect of pausing to the perception of (non-linguistic) disfluency, thus resulting in a weaker impact on the semantic judgment of negation scope.

Hypothesis 2: Intonation contour. The second hypothesis was widely confirmed as well. The shape of the intonation contour (or tune) turned out to be the most salient (and highly significant) cue for resolving scope ambiguities in German (Experiment I, Fig. 3B): As expected, a sequence of two pointed hats triggered the interpretation of a narrow scope whereas a flat hat pattern rather induced a wide scope interpretation. In Experiment II, the factor *intonation contour* reached significance as well but appeared to be a much weaker cue for the perception of a phrasal break (Fig. 4B), adding to the strong effect of pausing in the interaction. We interpret this finding in terms of cohesion between the two clauses: flat hats seem to enhance the cohesion between two information units whereas pointed hats are more likely to mark them as separate.

Hypothesis 3: Peak alignment. The first part of the hypothesis could not be confirmed, since variation in peak alignment did not have an influence on the perception of a phrasal break (Experiment II, Fig. 4C). However, our results were in line with the second part of the hypothesis: alignment differences did prove to have a significant influence on scope interpretations (Experiment I, Fig. 3C). As assumed, early peak accents resulted in an increased number of narrow scope readings whereas late peak accents induced more wide scope interpretations. This result was true both for the accent on the negation particle in the *main clause and* for the accent on the complement in the *subordinate clause*, which contradicts our (minor) hypothesis suggesting a different distribution of early and late peaks in the subordinate clause. Medial peak accents took a position between early and late peak accents for Düsseldorf and Viennese listeners, as expected, but evoked more wide scope interpretations in subjects from Kiel. Furthermore, we did not find an interaction between the factors *alignment* and *contour* that was relevant for scope interpretation, i.e. the position of the accent peak did not interact with the type of hat pattern presented in both clauses.

In fact, the results confirm the assumption that the type of accent in the subordinate clause is not decisive for a listener's scope interpretation. As can be derived from Fig. 3C, early peaks in the main clause had a stronger (or primary) effect on the subjects' decision in favour of a wide scope interpretation than early peaks in the subordinate clause.[13]

Against expectation, an early peak accent on the negation in the main clause did not lead to the perception of a subsequent phrasal break. This result supports the idea that the (lack of) cohesion between two clauses can also be indicated by other means than phrasing. According to previous studies (Swerts et al., 1994; Wichmann et al., 2000), early peaks evoke a stronger impression of givenness or finality than medial or late peaks. Following these studies, it is the narrow scope reading that can be interpreted as 'more final' than the wide scope reading, since the latter implies some sort of continuation: 'She did not read the novel because it is in English . . . BUT for another reason.' This line of argument is supported by the interaction between *alignment* and *pause* in Experiment I, showing a systematic relation between early peak accents and silent pauses which both add to the impression of 'finality' and thus favour a narrow scope reading.

The insertion of a phrasal break, which was found in Experiment II to be predominantly triggered by a silent pause (plus final lengthening), adds another factor that is important for the interpretation of alignment differences: if a boundary is perceived after the main clause, the first accent (on *nicht* 'not') becomes nuclear. Since an F0 peak has been shown to be earlier in nuclear accents than in prenuclear ones (Silverman and Pierrehumbert, 1990, for English; Schepman et al., 2006, for Dutch; Mücke et al., 2006, for German; also confirmed by Braun's 2007 data), the accent on the negation will be interpreted as having a late peak (compensating for the expected peak alignment within nuclear accents). Therefore, we would expect more 'yes' answers (i.e. wide scope readings) in utterances containing a break. However, this could not be found. Probably, the cues marking the juncture, in particular the silent pause, were much stronger than the rather subtle alignment difference.

Hypothesis 4: Dialectal influence. The hypothesis was confirmed since it could be shown that (a) the three parameters investigated are relevant perceptual cues for phrasing (with the exception of alignment) and for resolving scope ambiguities in all three varieties and (b) that there are systematic differences across the varieties with respect to the perception of negation scope.

[13] The following order of alignment combinations could be observed, with increasing numbers of wide scope interpretations from left to right: early & early < early & late < late & early < late & late.

Most importantly, and confirming the central part of hypothesis 4, we found an interaction between *variety* and *peak alignment* in Experiment I (see Fig. 3C). The stronger tendency for Viennese listeners to interpret the stimuli as having narrow scope can thus be related to the difference in alignment between Vienna and the Northern German varieties (as discussed in section 1.2): Viennese speech has been shown in acoustic and articulatory studies (e.g. by Mücke and colleagues) to display later peaks than Düsseldorf (and Kiel) speech. For this reason – and as hypothesised above – medial (and late) peaks are interpreted by Viennese listeners as 'earlier' as by listeners from the other varieties, leading to more narrow scope interpretations in the Austrian German variety.

Furthermore, not only alignment differences between varieties affect scope interpretations but also differences in the perception of the global contour, which is reflected by an interaction between the factors *variety* and *intonation contour*. Again, we found more narrow scope readings reported by Viennese listeners, presumably due to a less clear interpretation of a flat hat as a marker of wide scope in comparison with subjects from the other two dialects (see Fig. 3B).

In fact, overall results of the scope of negation experiment revealed that narrow scope was the unmarked interpretation for all three varieties. The silent reading experiment mentioned in sections 2.3.1 and 2.3.2 clearly confirmed this general bias. The most likely explanation for this finding is that the syntax of the test sentence did not suggest a strong cohesion between the two clauses. In contrast, a wide scope of negation was the less obvious reading, often triggered by a late accent peak (in Experiment I), which seemed to be a somewhat marked realisation for subjects from Kiel and Düsseldorf. Interestingly, listeners from Kiel judged medial peaks (which are often regarded as the neutral accent realisation; e.g. Grice et al., 2005) rather consistently as indicating wide scope, particularly in a flat hat contour. As shown in a recent psycho-phonetic study (D'Imperio et al., 2010), Northern German listeners perceived a medial accent peak as late if it was followed by a plateau – which is exactly what happens in a flat hat contour. Thus, we can assume that the medial accent peaks may have led to the perception of a relatively late peak in our study and therefore increased the number of wide scope readings for the Kiel variety.

Following Braun (2007), we have to take into account the ''linguistic use of peak alignment'' (2007:964). That is, later peaks are common for Southern German varieties in a 'neutral', i.e. *non-contrastive*, setting which can be attributed to the narrow scope reading of the test sentences: 'She did not read the novel' is a self-contained piece of information – the reason ('because it is in English') only provides additional information.[14] This may serve as another explanation for a larger number of narrow scope interpretations in Viennese subjects. On the other hand, later peaks are common for Northern German varieties in a *contrastive focus* setting, i.e. Northerners delay their accent peaks much further than Southerners in order to enhance the prominence of a contrastively accented word.[15] This may explain why more listeners from Kiel and Düsseldorf chose wide scope interpretations: later peaks sounded more appropriate for Northern German subjects to mark wide scope readings, since they signalled a contrast ('She did NOT read the novel because it is in English ... BUT for another reason').[16]

The peak alignment discussion leads to the research question as to which alignment differences between the varieties should be expressed as different accent types, i.e. in terms of phonological categories. Since most perceptual differences due to *peak alignment* were found to be rather subtle (despite their statistical significance in Experiment I), the following discussion is primarily meant to provide suggestions for a closer look into possible categorical distinctions between the intonational systems of the three varieties under investigation. The results for the Kiel variety seem to suggest a two-way distinction between early (H+L*) and medial/late ((L+)H*/L*+H) accents (in the main clause) which is in line with Kohler's (1991) findings for Northern German. Recall, however, that early peak accents are only realised as H+L* if they occur in pointed hat contours. In a flat hat, the L* is missing, and the term 'early peak' merely indicates that the H target is reached in the pre-accentual syllable (see end of section 2.2). Thus, it has to be stated clearly that our 'early peaks' can only partly be equated with H+L* accents. For Düsseldorf, the data rather suggest a categorical distinction between early/medial peaks (H+L*/H*) and late peaks (L+H*/L*+H), a contrast which is phonetically less clear than between early and medial/late peaks. Even less straightforward is the question of a phonological accent type distinction for Viennese. Table 1 in section 1.2 above as well as Fig. 3C suggest a similar pattern as for Düsseldorf, i.e. medial peaks are interpreted in the same way as early peaks, both contrasting with late peaks. However, since late peaks are obviously interpreted as earlier by Viennese listeners, there may not be a categorical distinction here at all. That is, the data suggest that while in Kiel and Düsseldorf alignment differences (either between early and medial/late or between early/medial and late) may be

[14] We conclude that an accent on the negation particle (as used in our perception experiments) marks a neutral reading. Nevertheless, this accent could also be interpreted as contrastive (in fact, corrective) since in a broad focus structure it is usually the argument (here: *the novel*) that receives the accent (cf. Uhmann, 1991).

[15] This has been shown at least for prenuclear accents (Braun, 2007). The accent on the negation particle is in fact prenuclear unless there is a phrasal break after the main clause.

[16] Another reason for the tendency to prefer the (unmarked) narrow scope reading found in Viennese listeners could be that the segmental realisation of the test sentence was felt to be unnatural since it was produced by a Northern German speaker. If a stimulus is not consistent with a listener's expectation towards one reading or another, he or she will probably be apt to choose a less marked interpretation.

'phonologised' in the sense that they are used to mark a specific scope of negation, the alignment differences in Vienna do not appear to be classifiable.

If alignment differences are 'phonologised', they have reached the interface between prosody and meaning. However, our results show that for the kind of semantic-pragmatic meaning we investigated here, namely the resolution of scope of negation ambiguities, not only specific types of pitch accent are relevant but also the global tune. In fact, resolving scope ambiguities proved to be predominantly tune-based in German, since the shape of the intonation contour (flat hat versus a succession of pointed hats) turned out to be the most important factor. Only in combination with the pitch accents (as the local building blocks of the tunes) and their alignment properties, however, as well as pausing, the whole picture emerges.[17]

4. Conclusion

The perception experiments on three German varieties have shown that the interpretation of negation scope is not as strongly linked to prosodic phrasing as expected. Although certain prosodic cues (in particular pausing and intonation contour) have a significant influence on both types of judgment, the relevance of these cues varies depending on the given task. An important cue for one task (such as the presence of a silent pause for the phonetic phrasing task) may be less relevant for another (such as the semantic-pragmatic interpretation of negation scope). Thus, the question of how German listeners resolve scope of negation ambiguities cannot simply be attributed to the presence or absence of a phrasal break between a main clause and a subordinate clause. It rather depends on a more general impression of 'cohesion' or 'connectedness' between the two clauses, indicated by prosodic means.

The cohesion expressed by a (global) tune in combination with local peak alignment variation proved to be the decisive factors for listeners from all varieties in resolving scope ambiguities: flat hat contours and late peak alignment patterns were the most important cues for wide scope readings in German. By contrast, pointed hats with early peak accents (especially in the main clause of the test sentence *Sie hat den Roman NICHT gelesen* (*weil er auf ENGlisch ist*) ('She did not read the novel (because it is in English')')), were typical of narrow scope readings. Thus, resolving scope ambiguities proved to be predominantly tune-based in German.

For the perception of phrasal breaks, the parameter *pause* (i.e. silence) turned out to be most influential, followed by *intonation contour* (pointed hats supported the perception of a stronger break), whereas differences in peak alignment did not play a role.

Acoustic and articulatory studies have shown that Viennese German displays later peaks than Northern German speech. This difference is reflected by our data in that we found an increase in narrow scope readings in Viennese listeners, suggesting that they interpreted (late) peaks as earlier than listeners from Kiel and Düsseldorf did. Proof for this dialectal difference was found in a significant interaction between *variety* and *peak alignment*. Generally speaking, however, the experiments demonstrated that the three varieties only slightly differed in the way they made use of prosodic cues in the interpretation of ambiguous utterances, both in terms of phrasing and the scope of negation.

Acknowledgements

We would particularly like to thank Jonathan Harrington, who inspired our interest in this research project, as well as Martine Grice, Doris Mücke and Anne Hermes for insightful comments and discussions. We are also grateful to Sónia Frota and two anonymous rewievers for their comments on an earlier version of this paper, to Ernst Dombrowski, Markus Linnemann and – in particular – Timo Röttger for their invaluable help with the statistics, and to Gudrun Schuchmann for collecting the Viennese data. The German Research Council, DFG, funded this research within the project STRETTS (GR 1610/2).

Appendix

Boundary tone experiment (mentioned in section 1.1)

In this additional experiment, we tested the influence of the type of utterance-final boundary tone on scope interpretation. As in Experiments I and II, the stimuli were based on the test sentence *Sie hat den Roman NICHT gelesen, weil er auf ENGlisch ist* ('She did not read the novel because it is in English'). The stimuli were kept constant in terms of accent position (accents on negation particle and complement) and peak alignment (medial accents), while the

[17] Note that we did not find a significant interaction between *intonation contour* and *peak alignment*, though.

parameters *pause*, *intonation contour* and *utterance-final boundary tone* were carefully varied. Each parameter had two levels, resulting in a set of eight different stimuli (pause/no pause, flat hat/pointed hats, rising/falling boundary tone).

Subjects and method

Twenty native speakers of Standard German (3 m, 17 f), all of them students at Cologne University, took part in the experiment. They were aged between 20 and 32 and did not report any hearing or speaking disorders. We did not control for the dialectal background of the subjects. In fact, nearly all participants originated from North Rhine-Westphalia.

The set of test stimuli was repeated five times, pseudo-randomised and separated into 4 blocks of 10 stimuli each. In total, the number of experimental stimuli was 60: 8 stimuli $\times$ 5 repetitions plus 20 dummies, which were presented in two additional blocks, i.e. a training block and a finishing block. These additional blocks were not evaluated. The stimuli were interspersed by pauses of three seconds in which the subjects had to make their decisions. The data were presented over loudspeakers in a sound treated room and to small groups of subjects simultaneously. One test session lasted eight minutes.

As in Experiment I, the data were obtained in a forced-choice test, i.e. subjects were asked to decide for each stimulus whether they preferred a narrow scope reading ('no, she did not read the novel') or a wide scope reading ('yes, she read the novel, but for another reason'). Participants marked their choices on a questionnaire.

Results

Fig. 5 shows the proportional distribution of wide scope interpretations ('yes'-responses) for the three prosodic parameters investigated across all subjects.

For inferential statistics, we constructed a model with binomial error function using the scope judgments (wide vs. narrow) as a categorical dependent variable. We included the categorical factors *repetition* (five levels), presence or absence of *pause* (two levels), *intonation contour* with the levels 'flat hat' and 'pointed hats', and *boundary tone* with the levels 'rise' and 'fall' as fixed effects. Wald's *z*-test indicated significant main effects of *pause* ($p < 0.001$), *intonation contour* ($p < 0.001$) and *boundary tone* ($p < 0.001$). The final model (with fixed effects) did fit the data significantly better than a null model (without fixed effects) (LRT: $\chi^2(3) = 456.97$, $p < 0.001$).

The data show that all three prosodic cues influence the interpretation of negation scope in German (and thus confirm the results for *pause* and *intonation contour* of Experiment I, although the effect of *pause* was much smaller in the three varieties investigated). Interestingly, the choice of utterance-final boundary tone had the strongest effect, with a rising boundary tone indicating wide scope in about 80% of the cases. That is, the rising boundary tone suggests that there is more to come, namely the actual reason for not reading the novel: 'she did not read the novel for reason A…, but for reason B'.

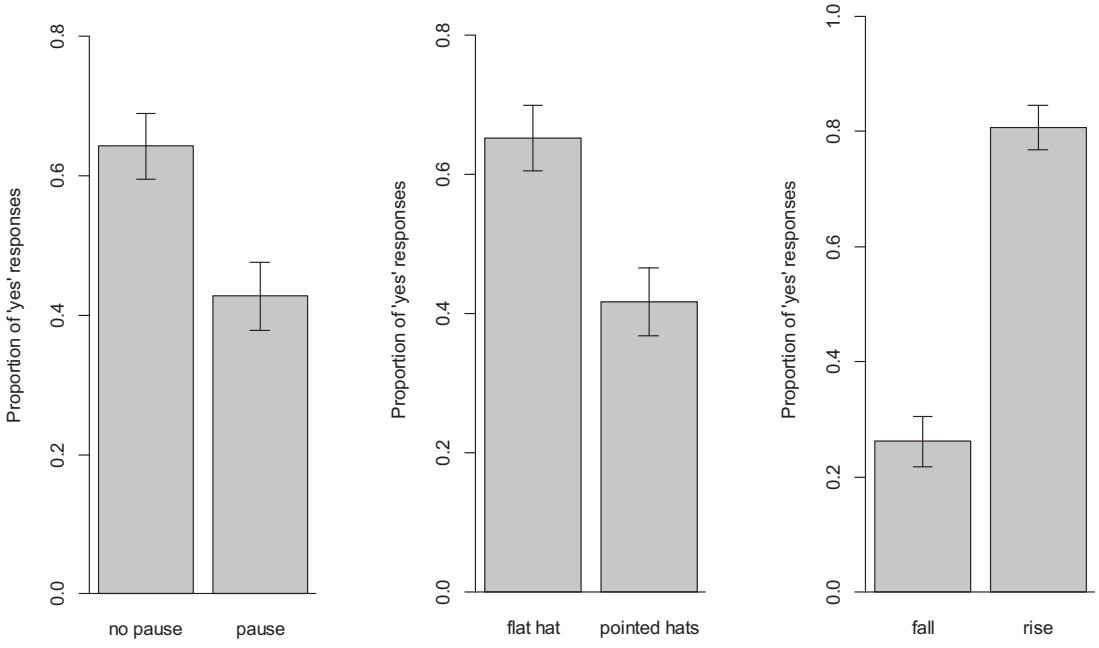


Fig. 5. Proportion means of 'yes'-responses for the factors *pause*, *intonation contour* and *boundary tone* (all subjects pooled, $n = 20$). Error bars represent 95% confidence intervals.

References

Adriaens, L.M.H., 1991. Ein Modell deutscher Intonation: eine experimentell-phonetische Untersuchung nach den perzeptiv relevanten Grundfrequenzänderungen in vorgelesenem Text. PhD thesis, University of Technology Eindhoven.

Atterer, M., Ladd, D.R., 2004. On the phonetics and phonology of ''Segmental Anchoring'' of F0: evidence from German. Journal of Phonetics 32, 177–197.

Avesani, C., Hirschberg, J., Prieto, P., 1995. The intonational disambiguation of potentially ambiguous utterances in English, Italian and Spanish. In: Proceedings 13th ICPhS, Stockholm, pp. 174–177.

Baayen, R.H., 2009. languageR: Data Sets and Functions With ''Analyzing Linguistic Data: A Practical Introduction to Statistics''. R Package Version 0.955.

Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language 59, 390–412.

Bates, D.M., Maechler, M., 2009. lme4: Linear Mixed-effects Models Using S4 Classes. R Package Version 0.999375-32.

Baumann, S., 2006. The intonation of givenness – evidence from German. Linguistische Arbeiten, vol. 508. Niemeyer, Tübingen.

Baumann, S., Grice, M., 2006. The intonation of accessibility. Journal of Pragmatics 38, 1636–1657.

Baumann, S., Becker, J., Grice, M., Mücke, D., 2007. Tonal and articulatory marking of focus in German. In: Proceedings 16th ICPhS, Saarbrücken, pp. 1029–1032.

Boersma, P., Weenink, D., 1996. PRAAT, A System for Doing Phonetics by Computer (Version 3.4). Report 132, Institute of Phonetic Sciences of the University of Amsterdam. URL: http://www.fon.hum.uva.nl/praat/.

Bolinger, D., 1965. Forms of English: Accent, Morpheme, Order. Harvard University Press, Cambridge, MA.

Bolinger, D., 1989. Intonation and Its Uses. Stanford University Press, Palo Alto.

Braun, B., 2006. Phonetics and phonology of thematic contrast in German. Language and Speech 49, 451–493.

Braun, B., 2007. Effects of dialect and context in the realisation of German prenuclear accents. In: Proceedings 16th ICPhS, Saarbrücken, pp. 961–964.

Braun, B., Ladd, D.R., 2003. Prosodic correlates of contrastive and non-contrastive themes in German. In: Proceedings 8th European Conference on Speech Communication and Technology, Geneva, pp. 789–792.

Crystal, D., 1992. An Encyclopedic Dictionary of Language and Languages. Blackwell, Oxford.

D'Imperio, M., 2001a. Focus and tonal structure in Neapolitan Italian. Speech Communication 33 (4) 339–356.

D'Imperio, M., 2001b. Tonal alignment, scaling and slope in Italian question and statement tunes. Proceedings of Eurospeech 99–102.

D'Imperio, M., House, D., 1997. Perception of questions and statements in Neapolitan Italian. Proceedings of Eurospeech 251–254.

D'Imperio, M., Gili Fivela, B., Niebuhr, O., 2010. Alignment perception of high intonational plateaux in Italian and German. In: Speech Prosody 2010, 100186. pp. 1–4.

Dainora, A., 2006. Modeling intonation in English: a probabilistic approach to phonological competence. In: Goldstein, L., Whalen, D.H., Best, C. (Eds.), Papers in Laboratory Phonology, vol. 8. Mouton de Gruyter, Berlin/New York, pp. 107–132.

De Zitter, A., 1992. The perceptual salience of melodic variation: contour shape and peak height. Journal of Phonetics 20, 181–188.

Féry, C., 1993. German Intonational Patterns. Niemeyer, Tübingen.

Floccia, C., Butler, J., Goslin, J., Ellis, L., 2009. Regional and foreign accent processing in English: can listeners adapt? Journal of Psycholinguistic Research 38 (4) 379–412.

Frota, S., 2002. Tonal association and target alignment in European Portuguese nuclear falls. In: Gussenhoven, C., Warner, N. (Eds.), Laboratory Phonology, vol. 7. Mouton de Gruyter, Berlin/New York, pp. 387–418.

Gibbon, D., 1998. Intonation in German. In: Hirst, D., Di Cristo, A. (Eds.), Intonation Systems. A Survey of 20 Languages. Cambridge University Press, Cambridge, pp. 78–95.

Grice, M., Baumann, S., 2002. Deutsche Intonation und GToBI. Linguistische Berichte 191, 267–298.

Grice, M., Harrington, J., 2003. Tonal realisation in two varieties of German. In: European Science Foundation Workshop: Experimental Approaches to Tone and Intonation, University of Oxford, April 2–5.

Grice, M., Baumann, S., Benzmüller, R., 2005. German intonation in autosegmental-metrical phonology. In: Jun, S. (Ed.), Prosodic Typology: The Phonology of Intonation and Phrasing. Oxford University Press, Oxford, pp. 55–83.

Gussenhoven, C., 1983. Testing the reality of focus domains. Language and Speech 26, 61–80.

Halliday, M.A.K., Hasan, R., 1976. Cohesion in English. Longman, London.

Hemforth, B., Koniecny, L., 2004. Scopal ambiguity preferences in German negated clauses. In: 26th Meeting of the Cognitive Science Society, Chicago.

Hirschberg, J., Avesani, C., 2000. Prosodic disambiguation in English and Italian. In: Botinis, A. (Ed.), Intonation. Kluwer, Dordrecht, pp. 87–95.

House, J., 2006. Constructing a context with intonation. Journal of Pragmatics 38, 1542–1558.

Jackendoff, R., 1972. Semantic Interpretation in Generative Grammar. MIT Press, Cambridge.

Jaeger, T.F., 2008. Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language 59, 434–446.

Kentner, G., 2010. Prosody and pragmatics in processing negative scope ambiguities. Poster presented at Phonetik und Phonologie, vol. 6, Goethe-Universität Frankfurt.

Kohler, K.J., 1991. Terminal Intonation Patterns in Single-accent Utterances of German: Phonetics, Phonology, and Semantics, vol. 25. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel, pp. 115–185.

Kohler, K.J., 1995. Einführung in die Phonetik des Deutschen. Grundlagen der Germanistik, vol. 20. Schmidt, Berlin, (2nd edition, 1st edition 1977).

Ladd, D.R., 2008. Intonational Phonology. Cambridge University Press, Cambridge, (2nd edition, 1st edition 1996).

Makarova, V., 2007. The effect of pitch peak alignment on sentence type identification in Russian. Language and Speech 50 (3) 385–422.

Mücke, D., Grice, M., Becker, J., Hermes, A., Baumann, S., 2006. Articulatory and acoustic correlates of prenuclear and nuclear accents. In: Proceedings Speech Prosody 2006, Dresden, pp. 297–300.

Mücke, D., Grice, M., Hermes, A., Becker, J., 2008. Prenuclear rises in Northern and Southern German. In: Proceedings Speech Prosody 2008, Campinas, pp. 245–248.

Mücke, D., Grice, M., Becker, J., Hermes, A., 2009. Sources of variation in tonal alignment: evidence from acoustic and kinematic data. Journal of Phonetics 37 (3) 321–338.

Niebuhr, O., 2003. Perceptual study of timing variables in F0 peaks. In: Proceedings 15th ICPhS, Barcelona, pp. 1225–1228.

Niebuhr, O., 2007. Perzeption und kognitive Verarbeitung der Sprechmelodie. Theoretische Grundlagen und empirische Untersuchungen. Language, Context, and Cognition, vol. VII. de Gruyter, Berlin/New York.

Peters, B., 2007. Form und Funktion prosodischer Grenzen im Gespräch. PhD thesis, Christian-Albrechts-Universität Kiel.

Peters, B., Kohler, K.J., Wesener, T., 2005. Melodische Satzakzentmuster in prosodischen Phrasen deutscher Spontansprache. Statistische Verteilung und sprachliche Funktion. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel 35a, 7–54.

Pierrehumbert, J., Hirschberg, J., 1990. The meaning of intonational contours in the interpretation of discourse. In: Cohen, P.R., Morgan, J., Pollack, M.E. (Eds.), Intentions in Communication. MIT Press, Cambridge, MA, pp. 271–311.

Pierrehumbert, J., Steele, S.A., 1989. Categories of tonal alignment in English. Phonetica 46, 181–196.

Price, P.J., Ostendorf, M., Shattuck-Hufnagel, S., Fong, C., 1991. The use of prosody in syntactic disambiguation. Journal of the Acoustical Society of America 90, 2956–2970.

Prieto, P., van Santen, J., Hirschberg, J., 1995. Tonal alignment patterns in Spanish. Journal of Phonetics 23, 429–451.

Prieto, P., D'Imperio, M., Gili Fivela, B., 2005. Pitch accent alignment in Romance: primary and secondary associations with metrical structure. Language and Speech 48 (4) 359–396.

Rathcke, T., 2006a. A Perceptual Study on *Russian* Questions and Statements, vol. 37. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel, pp. 51–62.

Rathcke, T., 2006b. Relevance of f0 peak shape and alignment for the perception of a functional contrast in Russian. In: Proceedings of the Conference on Speech Prosody, Dresden, Germany, pp. 65–68.

Redi, L., 2003. Categorical effects in the production of pitch contours in English. In: Proceedings 15th ICPhS, Barcelona, pp. 2921–2924.

Röhr, C., Baumann, S., 2011. Decoding information status by type and position of accent in German. In: Proceedings 17th ICPhS, Hong Kong, China, pp. 1706–1709.

Schaffranietz, B., 1997. Zur Unterscheidung und Funktion von restriktiven und appositiven Relativsätzen im Deutschen. Linguistische Berichte 169, 181–195.

Schepman, A., Lickley, R., Ladd, D.R., 2006. Effects of vowel length and ''right context'' on the alignment of Dutch nuclear accents. Journal of Phonetics 34, 1–28.

Seiler, H., 1960. Relativsatz, Attribut und Apposition. Harrassowitz, Wiesbaden.

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

Sievers, E., 1903. Über ein neues Hilfsmittel philologischer Kritik. Reprinted in Rhythmisch-Melodische Studien: Vorträge und Aufsätze (1912). Karl Winter, Heidelberg, pp. 78–111.

Silverman, K., Pierrehumbert, J., 1990. The timing of prenuclear high accents in English. In: Kingston, J., Beckman, M.E. (Eds.), Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech. Cambridge University Press, Cambridge, UK, pp. 72–106.

Swerts, M., Collier, R., Terken, J., 1994. On the prosodic prediction of discourse finality in spontaneous monologues. Speech Communication 15, 79–90.

't Hart, J., Collier, R., 1975. Integrating different levels of intonation analysis. Journal of Phonetics 3, 235–255.

't Hart, J., Collier, R., Cohen, A., 1990. A Perceptual Study of Intonation. Cambridge University Press, Cambridge.

Turk, A.E., 1999. Structural influences on boundary-related lengthening in English. In: Proceedings 14th ICPhS, San Francisco, pp. 237–240.

Uhmann, S., 1991. Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie. Niemeyer, Tübingen.

Ulbrich, C., 2002. A comparative study of intonation in three standard varieties of German. In: Proceedings Speech Prosody 2002, Aix-en-Provence, pp. 671–674.

Welby, P., 2003. Effects of pitch accent position, type and status on focus projection. Language and Speech 46, 53–81.

Wichmann, A., House, J., Rietveld, T., 2000. Discourse constraints on F0 peak timing in English. In: Botinis, A. (Ed.), Intonation. Kluwer, Dordrecht, pp. 163–182.

Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M., Price, P.J., 1992. Segmental durations in the vicinity of prosodic phrase boundaries. Journal of the Acoustical Society of America 91, 1707–1717.

Wunderlich, D., 1988. Der Ton macht die Melodie – Zur Phonologie der Intonation des Deutschen. In: Altmann, H. (Ed.), Intonationsforschungen. Niemeyer, Tübingen, pp. 1–40.