



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Phonetics 32 (2004) 177–197

Journal of
Phonetics

www.elsevier.com/locate/phonetics

On the phonetics and phonology of “segmental anchoring” of *F0*: evidence from German

Michaela Atterer¹, D. Robert Ladd*

Theoretical and Applied Linguistics, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Adam Ferguson Building, George Square, Edinburgh EH8 9LL, UK

Received 20 March 2002; received in revised form 10 June 2003; accepted 12 June 2003

Abstract

We measured the alignment of *F0* minima and maxima with segmental landmarks, in prenuclear rising accents in Northern and Southern German. As in earlier studies of other languages, we found consistent patterns of alignment. Both Northern and Southern German speakers align rises later than published data for Greek, English, and Dutch; Southern German speakers show later alignment than Northern speakers. The differences are small but significant. Moreover, native patterns of alignment are carried over into the German speakers' pronunciation of English. These findings argue against interpreting cross-language alignment differences in terms of distinct patterns of phonological association, and in favor of describing them in terms of quantitative phonetic realization rules.

© 2003 Elsevier Ltd. All rights reserved.

1. Introduction

A number of recent studies point to the conclusion that the temporal coordination of fundamental frequency (*F0*) and phonetic segments is highly lawful. Landmarks in the *F0* contour (such as local minima and maxima) are *aligned* in predictable ways with identifiable points in the segmental string (such as the onset of a stressed syllable). In this paper we demonstrate that the details of this “tonal alignment”, even in contexts that seem to involve the same general type of pitch accent, differ in consistent and specifiable ways from language to language and from one variety to another within the same language. This in turn has implications for the phonological representation of pitch accents.

*Corresponding author. Tel.: +44-131-650-3961; fax: +44-131-650-3962.

E-mail addresses: atterer@ims.uni-stuttgart.de (M. Atterer), bob@ling.ed.ac.uk (D.R. Ladd).

¹Institut für Maschinelle Sprachverarbeitung (IMS), Universität Stuttgart, Germany.

1.1. “Starred tones” and “segmental anchoring”

The starting point for modern phonetic research on tonal alignment is of course Bruce’s (1977) study of Swedish word accents, which introduced the idea that *F0* contours are best thought of as sequences of *F0* targets. Bruce’s great contribution was to realize that the *F0* features relevant to the Swedish word accent distinction are not necessarily those stretches of contour that happen to be delimited by the accented syllable, but rather features that can be defined and identified *independent* of the segmental material. Specifically, he proposed that the Swedish word accent is always manifested by a fall in pitch (a sequence of tonal targets H L) *in the general vicinity* of the accented syllable, and that the lexical accent distinction is based on the fact that the fall is aligned earlier, relative to the accented syllable, in Accent I than in Accent II. Bruce’s insight has since been applied to the study of languages in which the distinctions signalled by alignment are intonational rather than lexical (e.g., Kohler, 1987 on German, Verhoeven, 1994 on Dutch, D’Imperio, 2001 on Italian).

Despite the clear evidence for the phonetic relevance of alignment, however, its theoretical status in phonology has remained uncertain. Bruce himself seems to have regarded the word accent of Swedish as a single phonological type—a sequence of targets H L—and did not propose distinct phonological representations for Accent I and Accent II. It was only later, in Pierrehumbert’s (1980) influential dissertation on English intonation, that the notational distinction between “starred tones” and “unstarred tones” was introduced to deal with alignment distinctions. Specifically, Pierrehumbert notated the early-aligned rise of English as $L + H^*$ and the late-aligned rise as $L^* + H$. She proposed that in such contrastive pairs of bitonal accents one of the tones would be “starred” (aligned with the accented syllable) and the other tone would merely “lead” or “trail” (i.e., occur at some specified distance before or after the starred tone). This interpretation was subsequently adopted by Bruce (1987) for his Swedish word accent analysis: Accent I was notated as $H + L^*$ and Accent II as $H^* + L$.

The “star” notation caught on very rapidly, and quickly began to be used in ways that went beyond its original function of indicating the existence of alignment contrasts. First, it was widely assumed that in a bitonal accent exactly one of the two tones *must* be aligned with the accented syllable, i.e., a bitonal accent must be either $T_1^* + T_2$ or $T_1 + T_2^*$. This assumption seems to be related to the notion that the starred tone is metrically “stronger”: given the tenets of metrical phonology, only strong–weak and weak–strong are available as possible relations between the two phonological elements in a single metrical constituent. However, the internal metrical structure of pitch accents does not seem to have been subjected to serious theoretical consideration, except by Grice (1995).

Moreover, the idea that bitonal accents must exhibit one of only two patterns of alignment brings with it the implication that the notations $T_1^* + T_2$ and $T_1 + T_2^*$ could be interpretable even without any contrast of alignment. And indeed, the starred tone notation in practice rapidly took on what we might call a “crypto-phonetic” function. For example, several autosegmental descriptions of German intonation (e.g., Uhmman 1991; Féry, 1993; Gibbon, 1998; Grabe, 1998; Truckenbrodt, 2002) have used $L^* + H$ to indicate a rise in pitch that impressionistically begins rather late in the accented syllable and reaches its peak in the following syllable, even though no contrast with $L + H^*$ is posited. Similarly, the ToBI transcription system for English posits an accent type $H + ^1H^*$ without a corresponding $H^* + ^1H$. That is, notations like $L^* + H$ and $H + ^1H^*$

have acquired a use rather like IPA symbols, to refer impressionistically to phonetic details of alignment. This is reminiscent of the way in which both [d] and [t] may be used in transcribing Australian languages in which there is no voicing contrast, to convey impressionistic similarity to voiced and voiceless stops in languages that do have such a contrast.

The assumption that one of the two tones of a bitonal accent must be aligned with the accented syllable was explicitly called into question by Arvaniti, Ladd, & Mennen (1998, 2000), on the basis of their studies of the alignment of prenuclear accents in (Modern) Greek. In keeping with the assumptions about bitonal accents just discussed, Arvaniti et al. (1998) originally aimed to determine which of the two tones in a rising prenuclear accent in Greek is “the” starred tone. In fact, however, they found that both the beginning and the end of the accentual rise accompanying the stressed syllable of a non-phrase-final word are consistently aligned with what they called segmental “anchor points”, and that neither anchor point normally lies within the time spanned by the accented syllable itself. Specifically, the *F0* minimum at the beginning of the rise is aligned on average a few milliseconds before the onset of the stressed syllable, while the *F0* maximum is aligned on average 15–20 ms after the beginning of the following unstressed vowel. Arvaniti et al. (2000) therefore rejected the Pierrehumbert/ToBI notion of a single starred tone governing the alignment of bitonal pitch accents; instead, they suggested that both individual tones might be aligned quite precisely with specific points in the structure.²

Findings of “segmental anchoring” similar to those of Arvaniti et al. (1998) have since been obtained for other languages. Ladd et al. (1999) showed that accentual rises in British English are also consistently aligned relative to the segmental string. However, the details of alignment in British English are slightly different from Greek; specifically, the initial L, as in Greek, is aligned with the onset of the stressed syllable, but the H is aligned earlier, typically late in the immediately following consonant. In another study, Ladd, Mennen and Schepman (2000) showed that the alignment of prenuclear accentual rises in Dutch depends on whether the stressed vowel is phonologically long or short (and hence on the syllable membership of the following consonant). Specifically, the L is aligned (as in English and Greek) at the onset of the stressed syllable, but the H is aligned late in the stressed vowel when the vowel is phonologically long and midway in the following consonant when the vowel is phonologically short. Preliminary (unpublished) work by Ladd and Schepman suggests that a similar distinction based on vowel length is found in British English as well. As Ladd, Mennen and Schepman point out, this suggests rather strongly that the H is seeking to align with the “right edge” of the stressed syllable.

1.2. Implications of segmental anchoring for intonational description

One possible phonological way of thinking about segmental anchoring is in terms of the notion of “secondary association” proposed by Pierrehumbert and Beckman (1988). According to

²The original point made by Arvaniti et al. (1998) was that their findings support an “autosegmental” (or Bruce-style) analysis of intonation contours, in which the accentual rise is treated as a sequence of a low (L) tone and a high (H) tone associated with specific points in phonological structure, rather than a traditional description based on distinctive pitch movements. They note that the slope and the duration of the accentual rise (which are the obvious quantitative parameters for describing pitch movements) are variable; moreover, slope and duration are almost completely determined by the alignment and *F0* level of the L and the H, which are highly consistent and which are the obvious quantitative parameters for describing tonal targets.

Pierrehumbert and Beckman's general theory of tonal association, tones are underlyingly associated either to the *heads* or to the *edges* of the prosodic domains to which they apply. In addition, however, they can acquire "secondary associations" to other specific elements in prosodic structure. Pierrehumbert and Beckman developed the notion of secondary association in order to account for the phonetic behavior of initial high boundary tones in Japanese. Such boundary tones are assumed to be associated with the "left edge" of a prosodic word, but are aligned phonetically with a specific mora (usually the second mora of the word) depending on the phonological context. Secondary association thus serves as a phonological account of phonetic details of alignment.

The concept of secondary association has since been developed further by Gussenhoven (2000) and Grice, Ladd, and Arvaniti (2000), but was not taken up by Arvaniti et al. (2000) in their discussion of starred tones. Yet, it seems clear that Arvaniti et al.'s original suggestion—that individual tones might be *aligned* with specific points in the structure—can readily be recast in phonological terms by saying that the individual tones of a pitch accent may be *secondarily associated* with specific points in the structure. The pitch accent as a whole would be associated with the accented syllable that is the head of the domain to which it applies, but the individual tones that make up the pitch accent could have "secondary associations" to specific heads or edges within that domain. This approach is theoretically attractive because of the apparent fact that segmental anchoring seems to be targeting phonological "edges". If we take this approach, then cross-linguistic differences of phonetic detail could be treated as detailed differences of secondary association. For example, we might assume that in both Dutch and Greek the rising accent is a tonal sequence L + H, underlyingly associated with the accented syllable, but in Dutch the H tone is secondarily associated with the right edge of the accented syllable, whereas in Greek the H tone is associated with the following syllable.

However, it is not clear that fine differences of secondary association are the most appropriate way of describing cross-language differences of segmental anchoring. An alternative would be to assume that different languages could share a given pattern of phonological association, but realize it phonetically in different ways. For example, we might assume that in both Greek and Dutch the L and H tones of the rising accent are both associated with the accented syllable, but that the phonetic realization rules of Greek differ from those of Dutch in ways that are in principle storable. This approach would be in keeping with the notion of "language-specific phonetic rules" that has been developed by e.g., Pierrehumbert (1990).

We believe that empirical evidence can help choose between these two descriptive approaches to differences of segmental anchoring. If alignment patterns all appear to fall into a small number of cross-linguistic types, it would increase the plausibility of assuming that alignment differences can be analyzed in terms of secondary association. Indeed, it might turn out that what we have been calling the "crypto-phonetic" assumptions of the star notation are justified, i.e. that there are only two basic patterns of associating a bitonal accent with a syllable. If, on the other hand, we find an extensive range of subtly different alignment patterns, then we might more plausibly assume that the underlying association of tonal and segmental elements is the same in all of them, and that the difference involves language-specific phonetic realization rules or something of the sort. The latter case would be rather like what we find with voice onset time (VOT) in stops: any given language has at most a few phonologically distinct categories of VOT, but there is now good evidence (Cho & Ladefoged, 1999) that cross-linguistically VOT shows a continuum of values.

In order to discover whether the range of alignment patterns falls into a few cross-linguistic categories or suggests a continuum of values, we need new data from new languages. With this in mind, the goal of the present study is to extend the investigations of Arvaniti, Ladd, and their colleagues to two different varieties of Standard German. German was chosen because, although closely related to English and Dutch, it appears impressionistically to align accentual rises rather differently. As noted earlier, autosegmental analyses have tended to use $L^* + H$ to indicate rising accents in German, suggesting that the rises are aligned later than otherwise comparable rises in English. Moreover, within German, there is an impressionistic difference between Northern and Southern varieties that also seems likely to be based on alignment. This North–South difference has often been commented on (e.g. Sievers 1903), but never to our knowledge investigated systematically. Gibbon (1998: 93) explicitly mentions that at least Southern German is characterized by a “*right-displaced prominence peak*, that is, the syllable perceived as being accented has low pitch, and a pitch rise, often followed by a peak, occurs on one of the following syllables (ToBI $L^* + H\dots$)” [emphasis in original; note the “crypto-phonetic” use of the ToBI notation]. Our pilot studies³ suggested that German accents are indeed aligned later than what Ladd and his colleagues had found for English and Dutch, and that one difference between Northern and Southern German is indeed a matter of small phonetic differences in F_0 alignment.

We carried out two experiments. The first is based on controlled German speech materials, read aloud by both Northern and Southern speakers. Our goal was to produce basic alignment data on Northern and Southern German and to verify whether there is in fact an alignment difference between the two. The data from the first experiment can also be compared with data from studies on English and Dutch to see whether the impressionistic use of $L^* + H$ in annotating German might be based on facts of tonal alignment. In the second experiment, we recorded German speakers of English as a foreign language reading the English speech materials from the Ladd et al. (1999) study and comparing the results with the data from the native English speakers in the original study. This permitted us to investigate whether the difference between Northern and Southern German carries over into the pronunciation of English spoken by Germans, and made possible a direct statistical comparison of the alignment patterns of the different groups. Together, the two experiments shed light on the question of whether alignment patterns, cross-linguistically, fall into a small number of discrete patterns or show a continuum of values.

2. Experiment 1

The purpose of this experiment was to produce reliable data on the alignment of the beginning and the end of prenuclear accentual rises in German, and to compare it with the methodologically comparable data from English, Dutch and Greek. More specifically, we were seeking evidence for

³The pilot studies were carried out as course projects for the second author’s course on Prosody at Edinburgh University in 1999–2000. The study comparing English and German was performed by the first author, then a student on the MSc course in Cognitive Science, and the one comparing Northern and Southern German was performed by Daniela Heide, then a visiting undergraduate from the University of Wuppertal.

the following hypotheses:

- Alignment in German is later than in English and Dutch (i.e., in German, both the beginning and the end of the rise are aligned later with respect to the segmental structure of the stressed syllable).
- There is a difference in alignment between Northern German speakers and Southern German speakers, with the alignment in Southern German speech being later than that in Northern.

In formulating these hypotheses, we expected that the alignment of the accentual rise and the stressed syllable would exhibit the same segmental anchoring found in other languages, reviewed in Section 1.1. More fundamentally, we expected that in German as in the other languages, the first accented word would be accompanied by a sharp rise in *F0*, and we assumed that in some important sense these accentual rises are cross-linguistically comparable.

2.1. Method

The general method was that used in the studies by Arvaniti, Ladd and their colleagues discussed in the introduction. We designed speech materials to control relevant structural and phonetic variables, then had speakers of German read the materials aloud under laboratory recording conditions; in no case did we give the speakers any instructions about what sort of intonation to use. We then measured the acoustic variables of interest using an interactive waveform display package.

2.1.1. Speech materials

The list of German sentences read by the speakers consisted of 13 test sentences interspersed with 16 fillers. In each test sentence the phenomenon of interest was the rising *F0* movement accompanying the lexically stressed syllable (the “test syllable”) of the first content word (the “test word”). In order to keep the data as comparable as possible and to facilitate identification of segmental landmarks and *F0* minima and maxima, the test sentences all met the following criteria:

- The test word was either an adjective followed by a noun or a noun followed by a genitive construction. This normally ensured that a prenuclear rising accent was put on the test word followed by a nuclear accent on the following noun.
- The test syllable was preceded by one or two unstressed syllables, and followed by one or more unstressed syllables.
- The test syllable always contained a phonologically short (lax) vowel, in order to avoid any effect of vowel length of the sort found in Dutch by Ladd et al. (2000).
- The consonants of the test syllable were always sonorants (normally nasals, occasionally /l/), to ensure a smooth *F0* contour.

A full list of the test sentences is given in Appendix B.

2.1.2. Speakers

A total of 18 speakers of Standard German were recorded, nine Southerners and nine Northerners. The Southerners were all from Bavaria. The Northerners were all from the north-west of Germany (Schleswig-Holstein, Hamburg, Niedersachsen and the north of

Nordrhein-Westfalen), north of the bundle of isoglosses broadly dividing Northern dialects from Central and Southern ones. We do not enter here into detailed discussion of German dialectology (see e.g., Besch, Knoop, Putschke, & Wiegand, 1983; König, 1994), but simply note that the two groups of speakers come from widely separated areas and might be expected to exhibit differences in their pronunciation of Standard German. Indeed, a majority of the speakers can readily be assigned to the correct group on the basis of a few seconds' impressionistic listening, even by the second author, who is a fluent but non-native speaker of German.

The recordings of two speakers from each group were discarded before analysis. Reasons for this included too many mispronunciations of the English test words (in Experiment 2), reading too fast and monotonously (which would make identification of *F0* minima and maxima unreliable), and, in one case, too many years spent in foreign countries or other dialect areas. The seven Southern German speakers whose recordings were used had all lived in Bavaria all their lives, with only a few of them having spent 1 year in a foreign country. The seven Northern German subjects had all spent the last few months or years of their lives in Munich, where the recordings took place, but they had all lived at least the first 20 years of their lives in the Northern dialect area described above.

Of the usable recordings, each dialect group contained four male and three female speakers. The Southern speakers were aged between 22 and 44 years with an average age of 29 years. The Northern speakers were aged between 20 and 40 years with an average age of 27 years. The speakers are referred to in the results as NF1 (Northern female 1), SM2 (Southern male 2), etc. All speakers were either students or staff of the University of Munich.

2.1.3. *Recording and analysis procedures*

The recordings of the German speakers were made with professional equipment in the recording studio of the Institut für Phonetik und sprachliche Kommunikation of the University of Munich. The utterances were directly recorded on a Compaq Presario computer and digitized at 48kHz, using GoldWave software (Version 4.21).

As the materials for both experiments were recorded in a single session, speakers were told that they would have to read a number of German sentences and then a number of English sentences. Preliminary tests had shown that German speakers found some of the English words difficult to pronounce, so all speakers were given time to familiarize themselves with the English sentences before the recording began. They were asked to read each sentence naturally and at a normal speed, but again we note that no explicit instructions were given about, e.g., which words to emphasize or what sort of intonation to use. They read the German list first, followed by the English list. The recording session lasted about 10–15 min. Speakers were paid a small sum for their participation.

For analysis, the digitized utterances were transferred to a Sun workstation in the laboratory of the Department of Theoretical and Applied Linguistics at Edinburgh University. *F0* was extracted using the ESPS *get_f0* program with its default values of 10 ms frame shift and 7.5 ms correlation window. All the measurements were made interactively using ESPS Waves software, based on simultaneous visual displays of the waveform, wideband spectrogram and *F0* contour.

The following six landmarks were identified in each utterance:

- C0—the beginning of the initial consonant of the test syllable;
- V0—the beginning of the vowel of the test syllable;

- C1—the beginning of the final consonant of the test syllable;
- V1—the end of the final consonant of the test syllable, i.e., the beginning of the vowel of the following syllable;
- L—the beginning of the F_0 rise (local F_0 minimum); and
- H—the end of the F_0 rise (local F_0 maximum).

Fig. 1 shows an example of an actual test word contour and the labelling of the six landmarks.

The segmental landmarks were in general easy to identify from the spectrogram because of the spectral discontinuity at the moment of closure and the moment of release of the oral closure for the nasals. The F_0 minima and maxima were also generally easy to locate, in part because the use of sonorants in the test syllables ensured relatively smooth F_0 contours. There were only two types of problem cases common enough to warrant comment. The most common problem was the occurrence of minor F_0 discontinuities or “blips” at the closure or release of the nasals (these can be seen in Fig. 1). As far as possible these blips were ignored in determining H and L. The other occasional problem was that the apparent end of the accentual rise was not marked by a local F_0 maximum but only a change of slope, from steeply rising to slightly rising, in cases where the speaker produced an audible phrase boundary after the test word. In these cases the location of the change of slope was estimated by eye and marked as H. This can also be seen in Fig. 1.

Not every test item could be analyzed. The most common reason for discarding an item was the use of a syllabic nasal for the unstressed syllable following the test syllable, rather than a nasal-schwa-nasal sequence, in a word like *nonnenhafte*. This made it impossible to locate landmark V1 (the end of the consonant and the onset of the unstressed vowel). Occasionally, test sentences were discarded because the speaker used a completely different intonation contour from the one under investigation. An average of 11.6 utterances per speaker were usable for analysis (range 10–13).

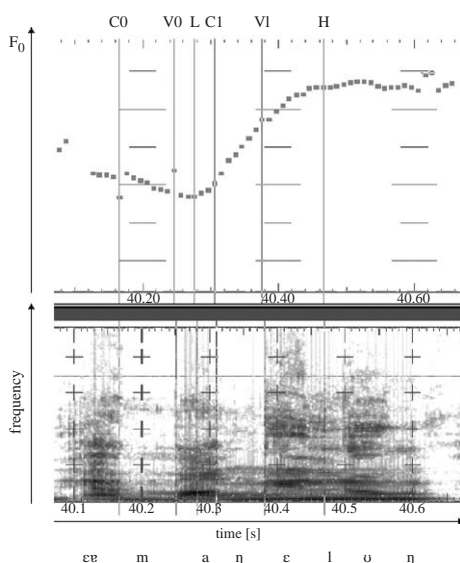


Fig. 1. Speaker SF1's utterance of the test word *Ermangelung*, showing the location of the four segmental and two F_0 labels. See text for more detail.

2.1.4. Reliability check

The measurements reported here were all made by the first author, but a reliability check was performed. The first author selected 28 items (two randomly chosen from each speaker, one from Experiment 1 and one from Experiment 2), which were then independently analyzed by the second author, and the analyses were compared. For each of the measured points (C0, V0, C1, V1, L and H) the mean absolute difference between the authors' labels was calculated. These were

- C0–6.1 ms,
- V0–4.5 ms,
- C1–5.4 ms,
- V1–5.4 ms,
- L–8.3 ms,
- H–4.2 ms.

While these absolute differences are small, they could affect our conclusions if they showed a consistent directional bias (i.e., if one author consistently located labels later than the other). However, there is no evidence for such a bias. For example, in the materials for Experiment 1, the second author's label for L was later in six out of 14 cases; in the materials for Experiment 2, the second author's label for L was later in eight out of 14 cases. Pooling across the two experiments and comparing by speaker dialect group, we found that the second author marked L later for exactly half of the 14 Northern utterances and exactly half of the 14 Southern utterances. There is thus no obvious reason to think that our conclusions are vitiated by measurement error or bias.

2.2. Results and discussion

2.2.1. Alignment in German and other languages

The basic data from Experiment 1 are reported in [Table 1](#). The table shows, for each speaker separately, the mean alignment of L and H. The alignment of L is reported in two ways, relative to both C0 (the “left edge” of the syllable) and V0 (the onset of the stressed vowel). The alignment of H is relative to V1 (the onset of the following unstressed vowel). The choice of segmental reference point is actually an important methodological issue, to which we return in [Appendix A](#).

As can be seen, all speakers align H within the following unstressed vowel. This means that for H, the German pattern of alignment is quite comparable to what [Arvaniti et al.](#) found for Greek, and rather later than the findings of [Ladd et al.](#) for British English and Dutch. The results for L are even more strikingly divergent from those of previous studies. Recall that [Arvaniti, Ladd and their colleagues](#) found that L is aligned at or slightly before the beginning of the onset consonant of the stressed syllable in Greek, British English, and Dutch; a similar finding is reported for Mexican Spanish by [Prieto, van Santen, and Hirschberg \(1995\)](#). In German, by contrast, speakers align L well within the initial consonant of the stressed syllable or even early in the stressed vowel. Since the methodology of the present study is effectively identical to that of these other studies, we can be fairly confident that the differences among these languages are genuine. Note also that similar L alignment in German is reported briefly by [Truckenbrodt \(2002\)](#).

A comparison between our results for the two groups of German speakers and the earlier results for Greek and British English is shown schematically in [Fig. 2](#). We do not attempt a direct statistical comparison, because the speech materials are not the same from language to language,

Table 1
Mean alignment data for Experiment 1

Speaker	H (ref V1)	L (ref V0)	L (ref C0)
NF1	30.1	−33.6	57.0
NM2	8.2	−40.2	41.5
NM3	26.6	−69.1	10.4
NM4	3.4	−30.7	49.3
NF5	26.3	−27.6	48.2
NM6	13.7	−50.4	19.8
NF7	41.8	−24.3	41.0
Northern grand mean	21.4	−39.4	38.2
SF1	27.5	9.4	79.3
SM2	58.8	6.3	75.0
SM3	41.9	−6.4	64.2
SM4	24.3	−4.4	57.8
SM5	28.3	−26.3	57.6
SF6	8.4	−8.7	56.8
SF7	48.5	8.6	84.3
Southern grand mean	34.0	−3.1	67.9

The columns show the distance in ms between an *F0* label (L or H) and a segmental landmark (C0, V0, or V1). A negative value indicates that the *F0* label occurs before the segmental label.

there are different numbers of speakers in different studies, typical segmental durations are different, and so on.

2.2.2. Alignment differences between Northern and Southern German

In order to test the hypothesis that there is a significant difference between the Southern and Northern patterns of alignment, we took the mean alignment values for each speaker (from Table 1) as the basis for one-way ANOVAs comparing Northern speakers and Southern speakers. We analyzed L and H separately. The difference between the two groups of speakers was highly significant for L. (For L measured relative to V0, $F(1, 12) = 22.70$, $p < 0.001$; for L measured relative to C0, $F(1, 12) = 14.80$, $p < 0.01$.) The difference for H, though in the same direction as the difference for L, did not reach significance ($p < 0.2$); possibly a larger group of speakers would show a clearer difference.

Though we cannot draw any firm conclusions for H, it seems clear that the alignment of the L—the beginning of the accentual rise—differs clearly and consistently between Northern and Southern German. In this connection it may also be worth noting that considering all the Northern utterances and all the Southern utterances as groups, the L occurs within the vowel in roughly only 10% of the Northern utterances, but in nearly two-thirds of the Southern ones.

2.2.3. Segment duration data

Though it is not strictly relevant to our central topic, we report a small but consistent difference between Northern and Southern speakers in the relative duration of the onset consonant and

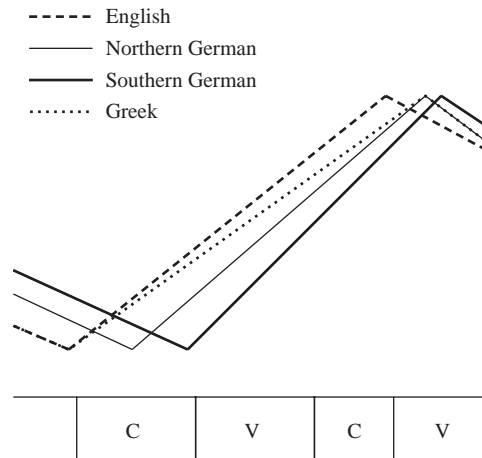


Fig. 2. Schematic representation of the alignment of rises relative to a stressed syllable in Northern and Southern German (based on Experiment 1), English (based on Ladd et al., 1999), and Greek (based on Arvaniti et al., 1998). Segment durations are necessarily somewhat idealized.

vowel of the stressed syllable. We observed that although the mean overall duration of the CVC sequence is virtually identical in the two groups of speakers, the *proportion* of the vowel duration is greater in the Southern speakers than in the Northern speakers. Specifically, in Southern speech the consonant preceding the stressed vowel is relatively shorter and the stressed vowel is relatively longer, compared to Northern speakers. There is no difference in the duration of the consonant following the stressed vowel, or in the overall duration of the CVC sequence as a whole.

Table 2 shows mean durations of both consonants and the vowel for each speaker separately, along with the duration ratio of the stressed vowel and the preceding consonant (“V/Cpre”). Fig. 3 graphically summarizes the North–South difference. In a one-way ANOVA carried out in the same way as for the alignment data above, the difference between the ratios is statistically significant ($F(1, 12) = 9.00, p < 0.02$).

It may be that in some way this difference contributes to the different auditory impression of Northern and Southern German intonation. Conceivably there is even a causal link between the later alignment and the relatively longer vowel, though at this point we have no way of knowing what that link might be. We report these data for completeness, in the hope that they may be of use to subsequent investigators.

3. Experiment 2

The purpose of Experiment 2 was to determine the extent to which the native alignment pattern of German speakers carries over into their pronunciation of English. Work by Jilka (2000) on native English speakers with advanced competence in German, and by Mennen (in press) on native Dutch speakers who speak Greek near-natively, suggests that native alignment patterns are likely to be carried over into an adult-acquired second language. By recording German speakers in English, we can directly compare German alignment with native English alignment on identical

Table 2
Mean segment duration data for Experiment 1

Speaker	Cpre	V	Cpost	V/Cpre
NF1	90.6	75.8	67.5	0.89
NM2	81.7	82.3	50.2	1.06
NM3	79.5	74.8	55.5	1.02
NM4	80.0	63.0	50.4	0.83
NF5	75.8	76.3	56.7	1.05
NM6	70.2	72.5	52.5	1.09
NF7	65.3	70.4	60.0	1.10
Northern grand mean	77.6	73.6	56.1	1.01
SF1	69.9	69.9	58.6	1.02
SM2	68.7	91.3	54.9	1.34
SM3	70.5	81.2	48.2	1.20
SM4	62.3	68.9	48.6	1.12
SM5	83.9	82.4	51.2	1.22
SF6	65.5	82.4	65.4	1.30
SF7	75.7	76.8	55.9	1.08
Southern grand mean	70.9	79.0	54.7	1.18

The columns show the durations of the consonant preceding the stressed vowel (Cpre), the stressed vowel (V), the consonant following the stressed vowel (Cpost), and the ratio of V to Cpre.

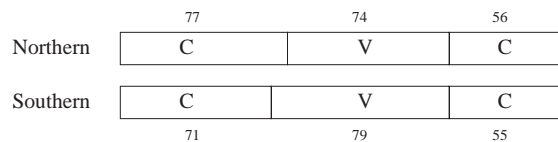


Fig. 3. Mean durations of the segments of the test syllables as spoken by Northern and Southern speakers.

speech material, and see whether later alignment is characteristic of a German accent in English. We can also compare Northern and Southern groups of German speakers to see whether the North–South difference demonstrated in Experiment 1 can be detected even when the speakers are speaking English.

3.1. Method

The basic method was the same as in Experiment 1, and in the previous work of Ladd, Arvaniti and their colleagues.

3.1.1. Speech materials

The speakers read a list of English sentences that consisted of 15 test sentences interspersed with 25 fillers. The sentences were a subset of the speech materials used in the study of alignment and

speaking rate in English by Ladd et al. (1999). This permitted direct statistical comparison with the results of English native speakers from that earlier study. A full list of the test sentences is given in Appendix B.

3.1.2. *Speakers*

The speakers were the same speakers as in Experiment 1. All studied English in school and all use it to one extent or another in their professional life. None could reasonably be considered “near native” and all but three speak English with either a readily identifiable German accent or an accent that is not necessarily obviously German, but markedly foreign.

3.1.3. *Recording and analysis procedures*

As noted above, the materials for Experiment 2 were recorded in the same session as those for Experiment 1. Analysis procedures were also the same as in Experiment 1. As in Experiment 1, some utterances were discarded before analysis; in Experiment 2 this often happened because the test word was mispronounced (e.g. *minister* rather than *minister*). An average of 12.6 sentences per speaker were analyzed (range 10–15).

3.1.4. *The native English data*

The English speakers analyzed by Ladd et al. (1999) had been recorded on professional equipment in the recording studio of the Department of Theoretical and Applied Linguistics at Edinburgh University. They were all undergraduates at the University of Edinburgh in their early twenties. They were speakers of Standard Southern British English. Eight speakers had been recorded, and two of them discarded because they were too monotonous. The remaining six speakers are referred to as FJ, FL, MG, MM, MN, and MR, where the first letters stand for female and male, respectively, and the second letter is an initial.

The English speakers were asked to read the English materials at three different rates, “normal”, “fast” and “slow”. The question of speaking rate was central to the Ladd et al. study. In principle it should be essentially irrelevant here, but because the average speaking rate of the German speakers of English was midway between the native English normal and slow rates, we compared the German data to English data at both rates. As one of the findings of the Ladd et al. study was that the alignment remains constant when speaking rate changes, it should of course make no difference whether we compare to the normal rate or the slow rate.

The original analysis had been carried out similarly to the analysis of the German speakers described above. In particular, the same interactive waveform display software was used as in Experiment 1, and the same *F0* points and segmental landmarks were identified.

3.2. *Results and discussion*

Mean alignment values are shown in Table 3 for L and H for Northern German and Southern German speakers, and for native English speakers at both the normal and slow speaking rates. The same values are reported as for the German materials in Experiment 1.

For each of the dependent variables in Table 3, we treated the data statistically as in Experiment 1, analyzing the effect of speaker group on the mean alignment values for each speaker in a series of one-way ANOVAs. As just noted, overall segment durations suggest that the average speaking

Table 3
Mean alignment data for Experiment 2

Speaker	H (ref V1)	L (ref V0)	L (ref C0)
<i>German speakers</i>			
NF1	1.9	-25.8	71.0
NM2	-2.8	-50.5	48.3
NM3	26.6	-60.0	37.0
NM4	-10.3	-48.4	38.1
NF5	33.0	-27.7	59.9
NM6	2.8	-42.5	41.3
NF7	23.7	-38.4	53.5
Northern grand mean	10.7	-41.9	49.9
SF1	21.1	-31.5	64.2
SM2	51.2	-24.8	55.9
SM3	28.3	-28.3	46.7
SM4	24.1	-6.7	71.4
SM5	11.4	-46.8	58.8
SF6	16.6	-21.6	60.1
SF7	29.8	-25.9	64.6
Southern grand mean	26.1	-26.5	60.2
German grand mean	18.4	-34.2	55.1
<i>English speakers</i>			
FJ	14.9	-87.1	-1.0
FL	-2.6	-72.4	-7.2
MG	-29.6	-105.1	-24.6
MM	-13.5	-83.6	0.8
MN	-18.3	-80.7	-1.3
MR	27.1	-92.5	0.1
Normal rate grand mean	-3.6	-86.9	-5.5
FJ	15.0	-75.4	9.2
FL	19.4	-89.2	-12.0
MG	-26.9	-126.1	-11.8
MM	-4.1	-114.0	-14.0
MN	-10.2	-94.7	-4.6
MR	117.4	-157.3	-7.3
Slow rate grand mean	18.4	-109.5	-6.8

The columns show the distance in ms between an *F0* label (L or H) and a segmental landmark (C0, V0, or V1). A negative value indicates that the *F0* label occurs before the segmental label. The data for the English speakers are taken from Ladd et al. (1999).

rate of the German speakers was midway between the native English normal rate and the native English slow rate, so we carried out the analysis separately for the normal rate data and the slow rate data from English.

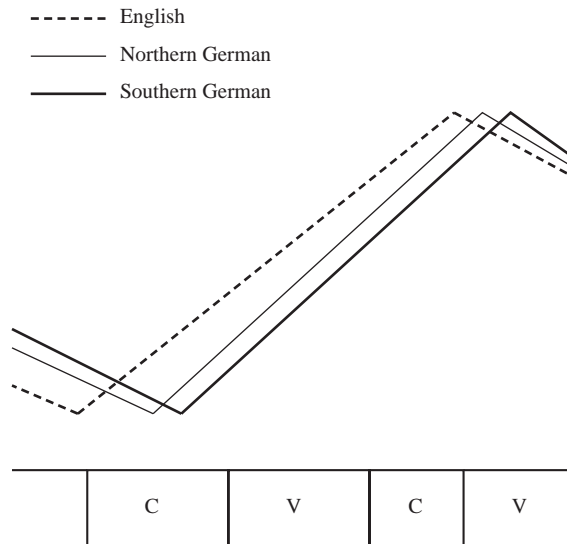


Fig. 4. Schematic representation of the alignment of the rises relative to the stressed syllable of the English test words for speakers of Northern and Southern German (based on Experiment 2) and English (based on Ladd et al., 1999). Segment durations are necessarily idealized.

We first compared the German group as a whole to the English speakers. For the normal rate, the effect of speaker group was statistically significant for all three dependent variables (for L relative to C0, $F(1, 18) = 128.67$, $p < 0.001$; for L relative to V0, $F(1, 18) = 65.03$, $p < 0.001$; for H, $F(1, 18) = 6.37$, $p < 0.05$). For the slow rate, the results for L were essentially the same, but the effect on H was not significant. This difference between rates can probably be attributed to the greater variability of H alignment in some of the English speakers at the slow rate, especially MR; when the same analysis is performed excluding MR, the result for H is significant as well ($F(1, 17) = 5.00$, $p < 0.05$).

We then compared the two groups of German speakers to each other. They differed significantly only for L relative to V0 ($F(1, 12) = 5.60$, $p < 0.05$), but the difference approached significance ($0.05 < p < 0.10$) for H and for L relative to C0.

Summarizing, the results of Experiment 2 show that, on identical English speech material, the German speakers align the accentual rise later than the native English speakers. This conclusion seems statistically robust for the alignment of the L and plausible for the alignment of the H. Experiment 2 further suggests (though this conclusion is less well supported) that the Southern German speakers align later than the Northerners, i.e., that the alignment patterns of their native variety of German are carried over into their pronunciation of English. The overall situation is shown graphically in Fig. 4.

4. General discussion

The results of our two experiments taken together appear to demonstrate that German accentual rises are aligned later than those in English or Dutch, and confirm past impressionistic

phonetic claims that Southern German rises are aligned later than Northern German ones. Moreover, our results suggest that German alignment patterns are carried over to the pronunciation of English as spoken by native speakers of German—indeed, they are carried over faithfully enough that the difference between the groups of Northern and Southern German speakers can be detected in the English data. These findings obviously reinforce the general claim of Arvaniti, Ladd and their colleagues that *F0* landmarks are consistently aligned with segmental anchors, and that patterns of alignment can vary from language to language (and, as it now appears, from one variety of a language to another.)

However, the most striking thing about the differences in alignment reported here is that, though apparently robust, they are rather small. This makes it rather difficult to treat them as based on differences in phonological representation. If we wanted to say that these differences in alignment reflect phonological differences of association, we would presumably have to say that the L in Dutch and English is associated with the left periphery of the accented syllable, in Northern German with the onset consonant of the accented syllable, and in Southern German with the left periphery of the accented vowel. This proliferation of docking sites for tones seems at best hard to justify theoretically. The difference in H alignment, assuming it is genuine, is even more problematical: we would have to distinguish between alignment with the left edge of the following unstressed vowel (Northern) and alignment with the middle of the following unstressed vowel (Southern), and explain why the H in the Northern variety is actually phonetically aligned *after* the left edge of the vowel. On the other hand, if the difference in H alignment is not genuine, it would mean that H alignment exhibits a rather considerable amount of inter-speaker phonetic variability. It would be at least puzzling if one and the same phonetic dimension were both the basis of fine phonological distinctions (L alignment) and wide inter-speaker variability (H alignment).

A more important reason why our findings make it difficult to support a purely association-based account of alignment differences lies in the fact that such an account presupposes categorical differences between distinct phonological representations. On the basis of the Greek, English and Dutch findings alone, it is not impossible to imagine finely differentiated specifications of tonal association, along the lines of “associate L with left periphery of syllable and associate H with following syllable”. Given the German data, however, this kind of account—where one pattern of association is categorically distinct from another—seems untenable. Our findings offer no support for the general assumption in the literature that the German rising prenuclear accent “is” $L^* + H$, distinct from the H^* and $L + H^*$ accents normally used in transcribing English. Phonetically, the alignment pattern we find for Northern German is midway between those of Dutch/English and of Southern German; it is difficult to escape the conclusion that we are dealing with a phonetic continuum of alignment. As we noted in the introduction, this finding makes tonal alignment rather comparable to voice onset time in stop consonants. Just as any given language has at most a few categories of voice onset time, but cross-linguistic phonetic detail shows a continuum of values, so the languages compared here may have only a single category “rising prenuclear accent”, but a continuum of phonetic alignment in the realization of this category.

It would, of course, be premature to assume that there is only a single category of rising accent in any of these languages. Indeed, one objection raised to our work by an anonymous referee is that it fails to take account of language-specific contrasts of alignment. While we acknowledge the

potential importance of this criticism, we believe that there is a strong *prima facie* case that the accents we investigate here are comparable with those investigated in English, Dutch and Greek by Arvaniti, Ladd and their colleagues. Recall that no instructions about intonation are given in any of these experiments; in all four languages the speakers are confronted with similar sentences and treat the first major accent in similar ways.

Indirect support for the view that we are dealing with a single category comes from the data in Experiment 2. According to a variety of work on the acquisition of phonology in a second language, it is common for speakers of a given first language (L1) to transfer phonetic properties of their L1 to a second language (L2) in cases where they identify a phonological category of the L2 with a category of their L1 (Flege, 1987, 1995). For example, having identified French /u/ with English /u/, English L1 speakers are unlikely to acquire an accurate pronunciation of French /u/, whereas they are more likely to acquire an accurate pronunciation of French /y/, which is difficult to identify with any English sound. Assuming (with Mennen, *in press*) that this finding can be shown to apply to intonational distinctions as well as to segmental ones, it would suggest that speakers readily identify “accentual rise” or L+H accent at least across the West Germanic languages, which would make them likely to carry-over the phonetic details of their L1 accentual rise into their pronunciation of L2. This is what we have found with the German speakers of English in Experiment 2. There could still, of course, be more than one category of “accentual rise” in one or more of these languages, but the evidence presented here suggests that the ordinary accentual rise used spontaneously by speakers of English, Dutch, German and Greek on the first accented word of a read sentence is in some way “the same thing” cross-linguistically. Exploring the more general implications of that sort of cross-language identification is beyond the scope of this paper.⁴

It is also worth noting that our results may provide some evidence for treating the accentual rise as a single phonological event at some level of analysis, as has been argued by Xu (1998 and elsewhere). The strongest interpretation of “segmental anchoring” would be that the two tones of a bitonal accent are completely independent of each other, each associated with a separate docking site in the structure, which determines its alignment. However, one interpretation of our comparative data would be that the alignment of the L and the alignment of the H are not independently determined: the alignment differences between English, Northern German, and Southern German may be completely parallel for the beginning and the end of the rise. That is, it appears that Southern German aligns *both* L and H later than Northern German, which in turn aligns *both* L and H later than Dutch or English. This would suggest that the two tones of a bitonal accent are in some sense part of the same phonological “event”. One could also, of course, point to the fact that the results for L are statistically well supported and those for H are not, as the basis of an argument that L alignment *can* vary independent of H alignment. This would strengthen Arvaniti et al.’s original view. We think it is more likely that the alignment differences between Northern and Southern German are genuine even for H, and their statistical significance is masked by greater variability.

⁴We would note, however, in response to a comment by an anonymous referee, that what we say here is not intended to contradict the discussion in Pierrehumbert, Beckman and Ladd (2000, Section 6). Indeed, the discovery and documentation of differences of fine detail in similar phonological categories in different languages, which is the thrust of our work, seems to us to be precisely in keeping with Pierrehumbert et al.’s general point.

Finally, we draw attention to a methodological point made by our study, namely that for a great many prosodic phenomena, serious investigation of differences between languages and dialects must be based on instrumental data. Notational distinctions such as $L^* + H$ vs. $L + H^*$, though they may certainly be useful for preliminary impressionistic transcription, are incapable of representing the range of fine phonetic differences that can be discovered instrumentally. Since these fine differences are relevant to many current theoretical debates on intonational phonology, such debates cannot usefully be conducted solely on the basis of impressionistic data written down in a symbolic notation. The tools for the necessary acoustic analysis are now so accessible that there is no reason for phonologists to restrict themselves to past techniques, and to the implicit assumptions that often accompany them.

Acknowledgements

The work reported here was supported by the UK Economic and Research Council under grant no. R000-23-7447 to Edinburgh University. We are grateful to Klaus Härtl and to the Phonetics Laboratory at the University of Munich for making the German recordings possible, to Michael Bennett, Stefan Evert, Astrid Schepman, and Cathy Sotillo for assistance with the analysis, and to Esther Grabe, Amalia Arvaniti, Colin Philips, an anonymous referee, and the participants in the Arbeitsgruppe Prosodie at the meeting of the Deutsche Gesellschaft für Sprachwissenschaft in Mannheim in February 2002, for useful comments and discussion.

Appendix A. Choice of segmental reference point

In the text we have referred to the problem of choosing a reference point for quantitatively characterizing the alignment of a given *F0* point. This should be an important methodological issue in the literature on *F0*/segmental alignment, but it has actually been little discussed. Various interacting problems arise.

One obvious problem is the speaking rate. To describe a given *F0* target *F* as occurring *x* milliseconds before segmental landmark *S* may mean something different depending on whether the speaker's average segment duration is 50 or 80 ms. A possible solution to this problem is to quantify alignment in proportional terms (*F* is aligned at *x*% of the interval between *S*₁ and *S*₂), a solution that has been adopted by Silverman and Pierrehumbert (1990) among others. However, this then entails a choice of two segmental landmarks (i.e., a choice of domain whose endpoints serve as landmarks), and the theoretical basis for this choice is by no means clear. Should alignment be described proportional to a segment? a syllable? a word? Even if there is a clear theoretical basis, there may be statistical consequences of the choice: in general, the larger the domain chosen (i.e., the greater the interval between *S*₁ and *S*₂), the greater the variance in the dependent variable *x* will be (because the duration of the domain will vary for many more independent reasons, and because the interval between *F* and either *S*₁ or *S*₂ will represent a smaller and smaller proportion of the overall domain).

The question of greater variance over larger domains also arises when we compare *F0* targets that are aligned at or near a segment boundary (like Southern German *L*) with those that are

aligned in the middle of a segment (like Northern German L). In general, the interval between an *F0* target F and a nearby segmental landmark S will show less variance than that between an *F0* target F and a more distant segmental landmark S' (again because of the likelihood that there will be more independent variation in the duration of intervening segments, etc.). This is exactly the reason for the different statistical conclusions about the North–South difference reported in Section 3.2: when the alignment of L is expressed relative to a landmark that lies between the extremes of the range of data points (namely V0), the difference between the two groups is statistically significant, but when it is expressed relative to a landmark that lies beyond one end of the total range of data points (namely C0), the difference does not reach significance.

We do not propose any definite solutions to these problems, but simply raise the issue as one that will need to be solved in order for the phenomenon of *F0*/segmental alignment to be fully understood.

Appendix B. Speech materials (test syllable in boldface)

B.1. German test sentences

Die **Verlängerung** der Ausleihfrist ist leider nicht möglich.

Die **Ernennung** Meiers zum Minister wurde nicht von allen Parteimitgliedern begrüßt.

In **Ermangelung** eines Lehrers übernahm ein Student den Unterricht.

Der **Mummenschanz** der Kostüme, das ornamentale Dekor und die synthetische Farbigkeit rückten den Film in die Nähe eines Pop-Märchens.

Die **Lungentätigkeit** des Patienten mußte künstlich aufrecht erhalten werden.

Seine **mangelhaften** Leistungen erlaubten es ihm nicht vorzurücken.

Die **nonnenhafte** Kleidung steht ihr überhaupt nicht.

Die **Vermengung** der Fächer Medizingeschichte, Wissenschaftstheorie und Bioethik könnte zu einem globalen Niveauverlust führen.

Auf **Verlangen** von Herrn Müller haben wir unser Sortiment erweitert.

Ein **nennenswerter** Unterschied war nicht auszumachen.

Die **mollige** Dame bezauberte durch ihr Lächeln.

Die **Minnesänger** von Nürnberg waren sehr berühmt.

Bei **Längengrad** Null wird die Universalzeit bestimmt.

B.2. English test sentences

There was a **nominal** fee for his services.

There is **phenomenal** interest in the products.

She got a **unanimous** vote for the proposal.

They got an **anonymous** call from a witness.

He made a **lemony** sorbet for dessert that evening.⁵

⁵This sentence was discarded from the analysis in Ladd et al. (1999), but was used here in order to make up for the fact that some of the other test words (especially *anomalous*, *unanimous*, *minuscule*) frequently had to be discarded because of mispronunciation.

She's a **minister's** wife in the Home Counties.
 There was an **anomalous** reading in the data.
 There were **monogrammed** sheets in the hotel rooms.
 There is a **minuscule** chance of surviving a plane crash.
 I need a **monosyllabic** word for my crossword puzzle.
 They sentenced the **militant** splinter group to five years.
 You need a **mineral** and vitamin supplement to get well.
 They charge a **minimum** rate for the use of their phone lines.
 He took a **mineral** enriched supplement every morning.
 They showed a **minimal** interest in what he had to say.

References

- Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics*, 26, 3–25.
- Arvaniti, A., Ladd, D. R., & Mennen, I. (2000). What is a starred tone? Evidence from Greek. In M. Broe, & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V* (pp. 119–131). Cambridge: Cambridge University Press.
- Besch, W., Knoop, U., Putschke, W., & Wiegand, H.E. (Eds.) (1983). *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. Berlin: de Gruyter.
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Bruce, G. (1987). How floating is focal accent? In K. Gregersen, & H. Basbøll (Eds.), *Nordic prosody IV* (pp. 41–49). Odense: Odense University Press.
- Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, 27, 207–229.
- D'Imperio, M. (2001). Focus and tonal structure in Neapolitan Italian. *Speech Communication*, 33, 339–356.
- Féry, C. (1993). *German intonational patterns*. Tübingen: Niemeyer.
- Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47–65.
- Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language speech research* (pp. 233–277). Timonium MD: York Press.
- Gibbon, D. (1998). Intonation in German. In D. Hirst, & A. di Cristo (Eds.), *Intonation systems: A Survey of twenty languages* (pp. 78–95). Cambridge: Cambridge University Press.
- Grabe, E. (1998). *Comparative intonational phonology: English and German*. (Ph.D. thesis, University of Nijmegen and Max Planck Institute for Psycholinguistics). MPI Series in Psycholinguistics, no. 7.
- Grice, M. (1995). Leading tones and downstep in English. *Phonology*, 12, 183–233.
- Grice, M., Ladd, D. R., & Arvaniti, A. (2000). On the place of phrase accents in intonational phonology. *Phonology*, 17, 143–186.
- Gussenhoven, C. (2000). The boundary tones are coming: on the nonperipheral realization of boundary tones. In M. Broe, & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V* (pp. 132–151). Cambridge: Cambridge University Press.
- Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent*. (Ph.D. thesis, University of Stuttgart). Vol. 6, no. 3, Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung, University of Stuttgart.
- Kohler, K. J. (1987). Categorical pitch perception. In *Proceedings of the eleventh international congress of phonetic sciences*, Vol. 5, Tallinn (pp. 331–333).
- König, W. (1994). *Dtv-atlas deutsche sprache*, revised edition. Munich: Deutscher Taschenbuch Verlag.
- Ladd, D. R., Faulkner, D., Faulkner, H., & Schepman, A. (1999). Constant “segmental anchoring” of *F0* movements under changes in speech rate. *Journal of the Acoustical Society of America*, 106, 1543–1554.

- Ladd, D. R., Mennen, I., & Schepman, A. (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America*, 107, 2685–2696.
- Mennen, I. (in press) Bi-directional interference in the intonation of Dutch speakers of Modern Greek. *Journal of Phonetics*.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Ph.D. dissertation, MIT.
- Pierrehumbert, J. (1990). Phonological and phonetic representation. *Journal of Phonetics*, 18, 375–394.
- Pierrehumbert, J., & Beckman, M. E. (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.
- Pierrehumbert, J., Beckman, M. E., & Ladd, D. R. (2000). Conceptual foundations of phonology as a laboratory science. In N. Burton-Roberts, P. Carr, & G. Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp. 273–303). Oxford: Oxford University Press.
- Prieto, P., van Santen, J., & Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23, 429–451.
- Sievers, E. (1903). *Über ein neues Hilfsmittel philologischer Kritik*. (Reprinted in: *Rhythmisch–Melodische Studien: Vorträge und Aufsätze*) (pp. 56–77). Heidelberg: Karl Winter, 1912.
- Silverman, K., & Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. In J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology I* (pp. 72–106). Cambridge: Cambridge University Press.
- Truckenbrodt, H. (2002). Upstep and embedded register levels. *Phonology*, 19, 77–120.
- Uhmann, S. (1991). *Fokusphonologie: Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie (Linguistische Arbeiten no. 252)*. Tübingen: Niemeyer.
- Verhoeven, J. (1994). The discrimination of pitch movement alignment in Dutch. *Journal of Phonetics*, 22, 65–85.
- Xu, Y. (1998). Consistency of tone–syllable alignment across different syllable structures and speaking rates. *Phonetica*, 55, 179–203.