Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Phonetics 33 (2005) 115–119
Journal of Phonetics
www.elsevier.com/locate/phonetics

Letter to the editor

Alignment of L and H in bitonal pitch accents: testing two hypotheses

Laura C. Dilley[a,*], D.R. Ladd[b], Astrid Schepman[c]

[a] *Speech Communications Group, Massachusetts Institute of Technology, Building 36, Rm. 549, 50 Vassar Street, Cambridge, MA 02139, USA*

[b] *Theoretical and Applied Linguistics, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Adam Ferguson Building, Edinburgh EH8 9LL, UK*

[c] *Division of Psychology, University of Abertay Dundee, Bell Street, Dundee DD1 1HG, UK*

Received 22 September 2003; received in revised form 26 November 2003; accepted 24 February 2004

Abstract

Two kinds of principles governing the timing or alignment of tonal targets have been proposed in modern theories of intonational phonology: first, targets may align with respect to the segmental string, and second, targets may align with respect to one another. We shed light on these two proposals by presenting an analysis of data from Ladd and Schepman (J. Phonetics 31 (2003) 81), which had found effects of syllable boundary location on the alignment of the fundamental frequency (F0) minimum at the beginning of a rising (L + H*) pitch accent in English. We investigated the effects of syllable boundary placement on the alignment of the accentual 1 maximum (H) relative to the F0 minimum (L) and relative to the onset of the stressed vowel (V). The syllable boundary manipulation significantly affects the duration of the interval between L and H, but not the duration of the interval between V and H. This suggests that the two tones in a bitonal L + H* pitch accent are aligned with respect to the segmental string, rather than each other. This contributes to a growing body of evidence that the fixed tonal alignment entailed in the original definition of bitonal pitch accents is at odds with phonetic facts.

1. Introduction

A growing body of cross-linguistic evidence suggests that the timing of fundamental frequency (F0) maxima and minima is important in signaling lexical and semantic distinctions and that such

*Corresponding author. Tel.: +1-617-253-1311; fax: +1-617-258-8304.
E-mail address: dilley@mit.edu (L.C. Dilley).

0095-4470/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2004.02.003

points are timed consistently with respect to the segmental string (Bruce, 1977; Kohler, 1987; Arvaniti, Ladd, & Mennen, 1998; Ladd, Mennen, & Schepman, 2000). Because theories differ in their predictions about the timing of such points, evidence concerning consistent alignment of F0 peaks and valleys is relevant to evaluating phonological models of intonation. (See Ladd, 2000 for a discussion.) This paper examines the timing of F0 maxima and minima in order to evaluate claims about the timing of tonal elements in bitonal L + H* pitch accents in English (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986).

In a recent paper in this journal, Ladd and Schepman (2003) presented evidence that the F0 minimum between two high accent peaks in English is the reflex of a phonological target (namely, a low tone) and not, as suggested by Pierrehumbert (1980), merely the lowest point of a "sagging transition" between the two peaks. A key aspect of Ladd and Schepman's argument is that the F0 minimum is consistently aligned with respect to the "left edge" or beginning of the second accented syllable, which was demonstrated by comparing the F0 contours on phrases like *It was Norman Elson* and *It was Norma Nelson.* The consistent alignment of the F0 minimum with respect to the beginning of the syllable is what one would expect of a bitonal (e.g. L + H*) accent, given earlier results such as those of Ladd, Faulkner, Faulkner, and Schepman (1999). Further evidence that the F0 minimum and following maximum form a unit (i.e. a bitonal accent) was provided by the fact that the scaling (i.e. F0 level) of the minimum was consistently related to that of the following F0 maximum but not the preceding F0 maximum. On the basis of this evidence, Ladd and Schepman argued that the F0 minimum represents a L tone in a bitonal L + H* pitch accent.

However, Ladd and Schepman did not report the alignment of the F0 maximum on the second accented word—the H* tone of the putative L + H* accent. It would be theoretically informative to determine whether this F0 peak occurs at a constant temporal interval after the F0 minimum, or whether the alignment of the maximum is determined independently of the alignment of the minimum. Current theories of the phonological structure of intonation suggest two different possible ways in which the timing of tonal elements might be controlled: either they may be aligned with respect to segmental landmarks such as syllable edges, or they may be aligned with respect to each other. Specifically, in a L + H* accent we might expect to find either that the L and the H are aligned with the segmental string independently of each other (e.g. at the beginning and end of the syllable), or we might expect the starred tone (H*) to be aligned to a fixed point in the stressed syllable and the other tone (L) to precede it by a fixed temporal interval. The first possibility was shown to be true of prenuclear rising accents in Greek by Arvaniti, Ladd, and Mennen (1998), and has since been used as the basis of valid predictions about alignment in other languages (e.g. Ladd et al., 1999 for English; Ladd et al., 2000 for Dutch). The second possibility was embodied in Pierrehumbert's early assumption that, in a bitonal pitch accent, one tone (the "starred tone") is aligned with a stressed syllable, while the other tone merely leads (or trails) the starred tone by a fixed temporal interval (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986). Little or no empirical evidence exists for the second possibility, yet the notion of a fixed relation between a "starred tone" and a leading or trailing tone persists; see Grice (1995) and Arvaniti, Ladd, and Mennen (2000) for discussion.

Ladd and Schepman's data provide an ideal testing ground for deciding between these two possibilities, a task which we have undertaken in the present study.

2. Method

Ladd and Schepman's specific experimental finding, based on two speakers, was that there is a significant effect of syllable boundary location on the alignment of the F0 minimum between the two accentual F0 peaks on names like *Norma Nelson* (''early boundary'') and *Norman Elson* (''late boundary''). When the syllable boundary was early, the F0 valley was on average 35 ms *before* the onset of the stressed vowel (e.g., the /ɛ/ of *Nelson*). When the syllable boundary was late, the F0 valley was on average 9 ms *after* the onset of the stressed vowel. On the basis of this finding we established two competing hypotheses regarding the alignment of the following F0 peak:

Constant interval hypothesis: If the tones in a bitonal pitch accent are separated by a constant temporal interval, then the manipulation of the syllable boundary location, which causes the L of a L + H* pitch accent to be later relative to the onset of the stressed vowel, should also cause the H to be later by the same amount. That is, there should be an effect of the syllable boundary manipulation on the temporal interval between the H and the onset of the stressed vowel, but no effect on the temporal interval between the H and the preceding L.

Segmental anchoring hypothesis: If the two targets in a L + H* pitch accent are aligned with respect to particular segmental positions or ''anchors'', then a manipulation which causes the L to be later in time should not affect the timing of the H. Instead, there should be an effect on the temporal interval between the L and the H. Specifically, the time interval between L and H should be longer when L is early (as in the early boundary condition) and shorter when L is later (as in the late boundary condition), but there should be no effect of boundary location on the alignment of the H relative to the onset of the stressed vowel.

These competing hypotheses were tested on the basis of Ladd and Schepman's data. We measured three variables: the time of the F0 minimum (L), the time of the following F0 maximum (H), and the time of the onset of the stressed vowel (V) that immediately follows the syllable boundary under manipulation. From these we calculated the time interval between the F0 minimum and following maximum (H–L) and between the F0 maximum and the onset of the vowel (H–V). All test names for which L, H, and V could be identified were examined ($n = 41$ and 45 for SF and EF, respectively).

If the Constant Interval Hypothesis is correct, there should be no significant effect of syllable boundary on H–L, whereas H–V should be significantly longer for the late boundary condition than for the early boundary condition. If the Segmental Anchoring Hypothesis is correct, H–L should be significantly shorter for the late boundary condition compared with the early boundary condition, and H–V should not be affected by the boundary manipulation.

We tested these hypotheses using two-tailed independent samples *t*-tests comparing the early vs. late boundary conditions. The tests were done separately for each of the two speakers.

3. Results and discussion

Results of the analyses are shown in Table 1. It can be seen that the mean value of H–L is significantly greater for both speakers when the syllable boundary is early than when it is late. In contrast, the mean value of H–V is not significantly different for either speaker when the early vs. late syllable boundary conditions are compared, although the difference does approach

Table 1
Duration differences and significance for time intervals between measurement points L, H, and V (in ms), comparing early and late syllable boundary conditions

	H–L				H–V			
Speaker	Mean (early)	Mean (late)	*t*	d*f*	Mean (early)	Mean (late)	*t*	d*f*
SF	138	99	3.264**	39	103	109	−0.392	39
EF	119	101	2.066*	43	64	78	−1.634	43

**$p<0.01$, *$p<0.05$.

significance for speaker EF. That is, as predicted by the Segmental Anchoring Hypothesis, the time interval from the F0 valley to the F0 peak (H–L) is significantly shorter when the syllable boundary is late, and the alignment of the peak relative to the stressed vowel onset is unaffected by the location of the boundary. By contrast, an effect on the time interval between the stressed vowel onset and the F0 peak (H–V), which is predicted by the Constant Interval Hypothesis, fails to appear.

These results support the conclusion that the two tones in a bitonal L + H* pitch accent are independently aligned with respect to the segmental string, and not with respect to each other. They contribute to a growing body of evidence that the alignment of tonal targets is specified relative to segmental positions. They also represent a clear failure of the idea that there are fixed time intervals between tones or more generally that speech F0 movements have fixed durations. As stressed by Arvaniti et al. (2000), the accumulating evidence that phonetic facts about alignment are not adequately captured in current intonational models highlights the need for revisions to prosodic theory.

Acknowledgements

We thank Mary Beckman for asking us the question that led us to this analysis. We also thank Doug Whalen and two anonymous reviewers for helpful feedback on an earlier draft.

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.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook*, *3*, 255–309.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerups.

Grice, M. (1995). Leading tones and downstep in English. *Phonology*, *12*, 183–233.

Kohler, K. (1987). Categorial pitch perception. In U. Viks (Ed.), *Proceedings of the 11th international congress of phonetic sciences*, Tallinn, *5*, 331–333.

Ladd, D. R. (2000). Tones and turning points: Bruce, Pierrehumbert, and the elements of intonational phonology. In M. Horne (Ed.), *Prosody: Theory and experiment—studies presented to Gösta Bruce* (pp. 37–50). Dordrecht: Kluwer.

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.

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics*, *31*, 81–112.

Pierrehumbert, J.B. (1980). *The phonology and phonetics of English intonation*. Ph.D. thesis, MIT.