Semantic judgments as evidence for the intonational structure of Dutch

Carlos Gussenhoven

Department of Linguistics, Radboud University Nijmegen
School of Languages, Linguistics and Film, Queen Mary, University of London
c.gussenhoven@let.ru.nl

Abstract

A perception task elicited judgments from 20 Dutch listeners involving a choice between a modal and a lexical meaning of three ambiguous words in Dutch. On the basis of descriptions in the literarture, it was hypothesized that the lexical interpretation is favored if (a) there is an intonational boundary between the ambiguous word and the following pitch accent, and (b) if pitch accents on either side of the boundary are identical. The results support this assumption. The data provide semantic evidence for a number of positions in the literature. First, the accentual rise-fall must be a HL contour (i.e. a fall, H*L) rather than a LH contour (i.e. a rise, LH*). Second, the results indicate that the presence of a low target immediately after the H-marked accented syllable signals an intonational boundary if the pitch remains low level until the next pitch accent, but not if the pitch after the post-accent low rises to the next accent. These results favor an analysis as given in the British English literature for English and as provided by ToDI for Dutch, and fail to support the contrary analytical assumptions made for MAE ToBI.

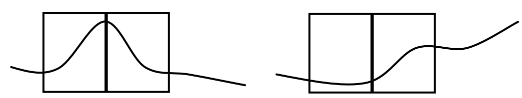
1. Introduction

Generally, morpho-syntactic structures do not come with dedicated intonational structures [1] (p. 67). Yet, particular constructions do sometimes appear to have specific default intonation patterns. For instance, wile the distinction between restrictive and non-restrictive constructions can sometimes be hard to draw [1, p. 194], in the typical non-restrictive case an appositive and the constituent it modifies are separated by a phrase boundary and have identical pitch accents, illustrated in (1) and (2) cf. [2, p. 85]. The most likely interpretation of these expressions is one whereby the speaker has only one brother, who is called John. A structure like (3), where the two constituents have different pitch accents, may be ambiguous between the above non-restrictive interpretation and a restrictive interpretation, whereby the speaker has more than one brother and *John* serves as a defining addition. The same is true for (4), where there is no intonational boundary between the constituents (transcriptions are in ToDI [3]).

```
My brother      John            (1)
      |           |
%L  H*L  %%L H*L   L%

My brother      John            (2)
      |           |
%L   L*H  %%L L*H   H%

My brother      John            (3)
      |           |
%L  L*H  %%L H*L   L%

My brother    John              (4)
      |         |
%L  H*L       H*L  L%
```

The above observations apply to a range of constructions. For instance, sentence-internal modal adverbs like *only* are typically not separated from the following constituent, as in *That punishment seems too harsh. She was ONly a little LATE*. Conversely, an predicative adjective like *alone* would typically be separated form the following place adverbial by a boundary *and* have the same pitch accent as that adverbial, as in *She was aLONE in the PUB*. If all accented syllables have H*L-pitch accents, the difference between the presence and the absence of the boundary comes out in the alignment and the depth of the valley between the peaks. That between -*LONE* and *PUB* is both earlier and lower than that between *ON*- and *LATE*. In Dutch, the equivalents of *only* and *alone* are identical, *alleen*, and it is this ambiguity between the modal and lexical meanings which the experiment reported here was aimed to exploit.

The intonation system of English being complex, it is not surprising that there are competing phonological analyses, even within the Autosegmental-Metrical tradition. The best-known analysis is that by Janet Pierrehumbert and Mary Beckman, in particular its modified version MAE ToBI [4]. ToDI, intended for Dutch, but virtually equally applicable to English (cf. chapter 15 in [5]), differs from MAE ToBI in a number of respects, most notably in having an off-ramp analysis of pitch accents, following the usual analyses in the pre-autosegmental descriptions of British English (e.g. [6], [2]). In an off-ramp analysis, the striped portions of the contours in Fig. 2 are taken to reflect the phonological elements. MAE ToBI is an on-ramp analysis, in which the dotted portions are assumed to reflect the phonological element.



MAE ToBI	L+H*	L-	L%		L*	H-	H%
ToDI	%L	H*L	L%	%L	L*H		H%

Figure 1. *On-ramp (dotted section and first line) and off-ramp (striped section and second line) analyses of nuclear falling and rising f_0 contours.*

The decision between these approaches has consequences for the rest of the analysis. Relative to an on-ramp analysis, an off-ramp analysis typically has one less data point to account for after the pitch accent, which will affect the requirements for the phrasing structure. An on-ramp analysis thus assumes two independent L-tones and H-tones (L-L% and H-H%) after the pitch accent, and two intonational constituents in which these tones are final boundary tones, the intermediate phrase (ip) and the intonational phrase (IP). Second, an off-ramp analysis of a rise-fall leaves the presence of a boundary after a fall undecided, while in an on-ramp analysis of a rise-fall the post-peak target must signal a boundary. Third, identity of an accentual peak is defined by the rise (L+H*) in an on-ramp analysis, but by the fall (H*L) in an off-ramp analysis.

In the perception experiment reported here these questions are addressed in a task exploiting the interpretative preferences of appositive-like structures in Dutch.

2. Method

2.1. Materials

The experiment involved a semantic task in which subjects had to choose between a modal and a lexical interpretation of a three ambiguous Dutch words, *alleen* ('just' and 'alone'), *gewoon* ('just' and 'normal'), and *vast* ('surely' and 'stuck'). The words appeared in three short and three long carrier sentences which allowed for both modal and lexical interpretations. The ambiguous words appeared in non-final position. In the short condition, a potential phrase boundary appeared immediately after the ambiguous word. In the long condition, at least one unaccented syllable appeared before the boundary. The sentences are given in (5). The a-gloss gives the lexical interpretation and the b-gloss the modal one. The bracketed words represent the long condition.

(5) - Hij zit alLEEN (met die man) in het caFÉ
- (a) 'He is alone (with that man) in the pub'
- (b) 'The thing is, he's in the pub (with that man)'

- Doe maar geWOON (hier) zoals WILlem
- (a) 'Act normally (here), like William'
- (b) 'Simply act (here) like William'

- Ze zit VAST (met dat ding) op de SNELweg
- (a) 'She has got stuck (with that vehicle) on the motorway'
- (b) 'She must be on the motorway (with that vehicle)

The sentences were recorded by a 26-year-old female native speaker of Dutch and after being downsampled to 16kHz the recordings were manipulated in Praat such that all f_0 information was replaced with eleven pre-determined f_0 contours.

As there were 11 contours, the full stimulus set contained 66 stimuli, which were randomized once and then reversed. Two cd-roms were prepared, one for each order, with 5 randomly chosen stimuli inserted at the beginning and one at the end, whose scores were ignored. Each stimulus was presented twice in succession, with an interval of 600 ms. The ISI between stimulus pairs was approximately 4 s. Twenty native speakers of Dutch were recruited from the student population at Radboud University Nijmegen and paid a small fee for participating. Their answer sheets showed a 5-point scale for every trial, with non-ambiguous paraphrases of the two interpretations on either side of the scale. The two test versions had reversed orientations of all scales. Listeners were instructed to listen to each stimulus and pay special attention to the intonation. They were asked to indicate on each scale whether the meaning of the sentence they heard corresponded strongly (-2) or less strongly (-1) to the paraphrase on the left or (strongly (2) or less strongly (1) to the paraphrase on the right, or whether it was not possible to make a choice (0). Subsequently, the scores were transferred to a scale from 1 to 5, such that 1 corresponded to the most lexical interpretation and 5 to the most modal interpretation.

2.2. Comparison I: the effect of the boundary

Six of the artificial contours involved the presence of a boundary between identical pitch accents. Specifically, prenuclear contours H*L, H*LH and L*H were combined with nuclear contours H*L L%, H*L H% and LH H%, respectively, giving three two-accent contours. Subsequently, an internal boundary was introduced by replacing the prenuclear pitch accents with H*L L%, HL H% and LH %, respectively, following ToDI conventions [3]. This gave us 3 (pitch accents) × 3 (words) × 2 (boundaries) × 2 (lengths) or 36 stimuli. Fig. 2 gives the six artificial contours, with dotted sections representing the f_0 in the long condition. The speech files were not subjected to any other adjustments, and the six sets of stimuli were therefore identical with the exception of the manipulated f_0.

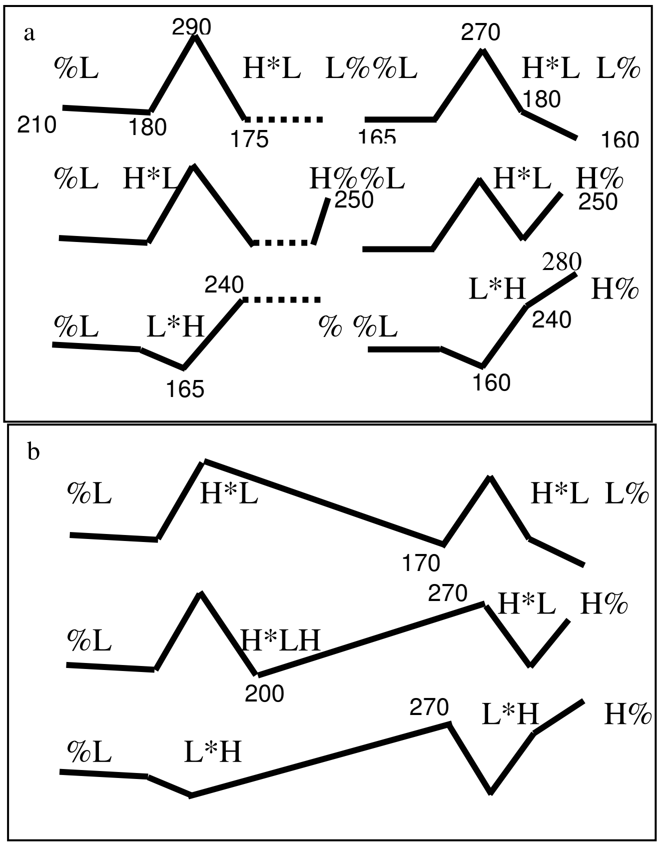


Figure 2. *Contours with two identical pitch accents with intervening intonational boundary (panel a) and without (panel b). Dotted sections are present in the long condition. f_0 values for identical targets are not repeated.*

A repeated measures MANOVA was run to test the null hypothesis that the intonational boundary does not affect the interpretation of the ambiguous words, with BOUNDARY (2 levels), WORD (3 levels), PITCH ACCENT (3 levels) and LENGTH (2 levels) as independent variables. For none of these factors was Mauchly's test for sphericity significant, and uncorrected significance levels are therefore reported. There was a significant threeway interaction WORD × PITCH ACCENT × LENGTH (F [4, 4.08] = 2.843, $p < .05$), but there was no other significant interaction. Main effects were found for all four variables: BOUNDARY (F [1, 32.51] = 14.55, $p < .01$), WORD (F [2, 9.83] = 5.83, $p < .01$), PITCH ACCENT (F [2, 22.05] = 6.214, $p < .01$) and LENGTH (F [1, 58.39] = 21.77, $p < .01$). I report the data for the long condition in Fig. 3 for the three pitch accents, for each word separately. As will be clear the prediction that the presence of the boundary favours the lexical interpretation is fulfilled in eight out of nine cases, by a constant margin of about half a scale point.

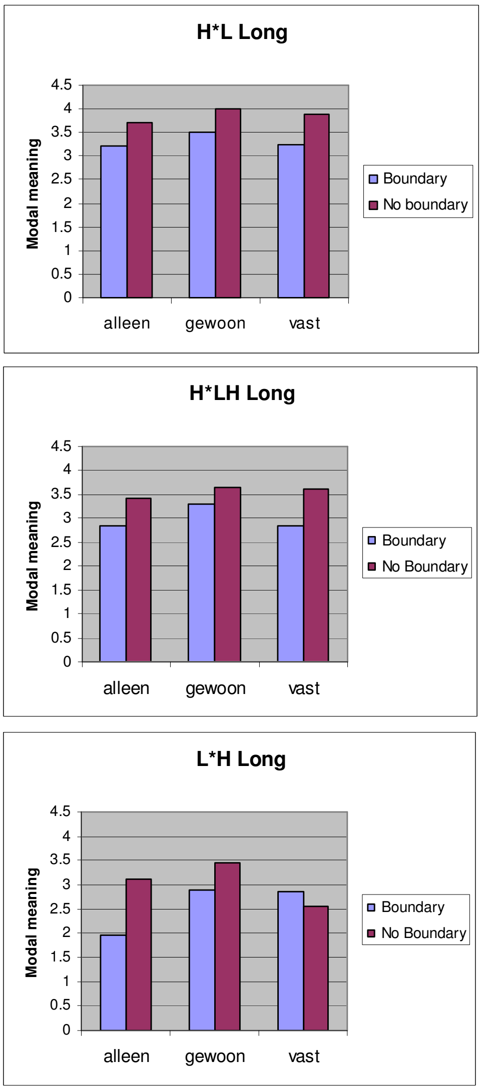


Figure 3. *Mean modal-scores for three contours and three ambiguous words separately in the long condition.*

The one exception concerns the word *vast* in combination with the %L L*H (% %L) L*H H% contour. This is explained by the incompatibility of the modal interpretation 'surely' with the interrogative intonation (cf. **?Is she surely on the motorway?*). Listeners must have rejected the modal interpretation for that reason. This interpretation is strenghthened by the much larger difference between the two boundary conditions than of the other eight differences in the opposite direction. The significant effect of LENGTH is due to the fact that, overall, the long condition attracts higher modal-scores than the short condition. In the short condition, three out of the eight cases that were significant in the long condition were not significant. This is probably due to the lesser salience of the boundary in the short condition, where the f_0 difference was in the second half of the preboundary syllable and the first post-boundary syllable(s). In this condition, the difference between absence and presence of the boundary was often very hard to hear in the short sentences (cf. Fig. 2).

To test the more specific prediction that low pitch immediately after a prefinal peak signals a boundary if the pitch rises to the new accent, but not if the pitch remains level until the next accent, we compared the scores for %L H*L L%%L H*L L% (cf. the top contour in panel a of Fig. 2, also known as the 'two pointed hats'), with those for %L H*LH H*L L%, i.e. the middle contour of panel (b) in Fig. without the final rise, the 'jackknife' of [6]. A repeated measures ANOVA was run with CONTOUR (2), WORD(3) and LENGTH (2) as independent variables. Mauchly's test for sphericity was not significant for any interaction or for WORD, and uncorrected significance values are reported. There were no significant interactions. Main effects were found for CONTOUR (F [1, 19] = 4.63, $p < .05$) and LENGTH (F [1, 19] = 28,80, $p < .001$). The hypothesis that there is no boundary inside the 'jackknife', but that the 'two pointed hats' are separated by a boundary was confirmed by the fact that the 'jackknife' attracted higher modal scores in the short condition (2.80 versus 2.70) and the long condition (3.98 versus 3.32).

2.3. Comparison II: the effect of pitch accent identity

A second interest was in the effect of the phonological identity of the pitch accents on either side of the boundary.

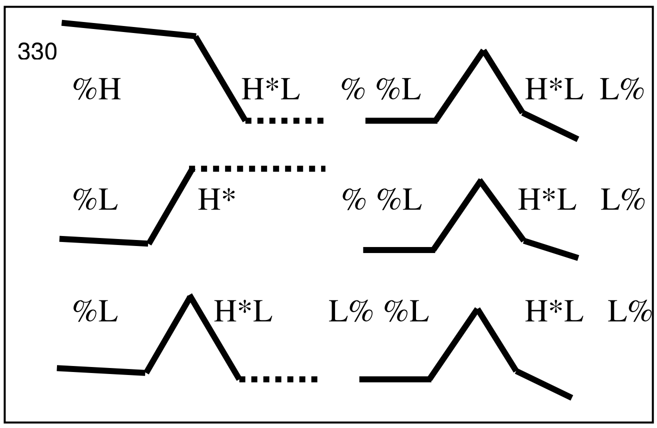
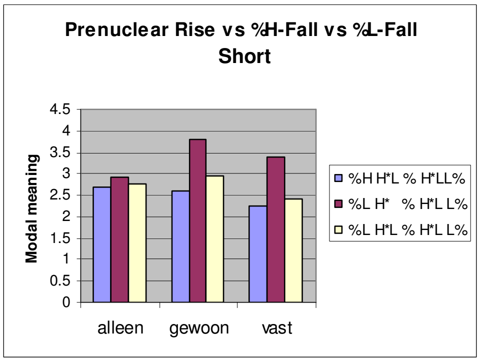


Figure 4. *Three contours with a fall (top), a rise (middle) and a rise-fall (bottom) preceding a rise-fall. f_0 values for targets that were given in Fig. 3 are not repeated.*

Our interest was in the question which half of the accentual rise-fall represented the element that defines that identity, the rise or the fall. To this end, we compared the scores for the %L H*L L%%L H*L L% contour, where pitch accent identity is phonetically unambiguous, with %H H*L L%%L H*L L%, where only the falling part of the first accentual peak is maintained, and %L H* %%L H*L L%, where only the rising part is maintained. The three contours are graphically represented in Fig. 4, with the base line contour, the 'two pointed hats' in third position.

Phonologically, the preservation of the fall leaves the presence of the H*L pitch accent unaffected, since the presence of the rise is determined by the choice of initial boundary tone, leaving the pitch accent unaffected. However, the preservation of the rise at the expense of the fall can only be achieved by changing pitch accent H*L into H*, assuming an off-ramp analysis, as in ToDI. The middle contour, where pitch accents are not identical, must thus be expected to favour the modal meaning relative to the other two contours, where pitch accents are identical. Figure 5 gives the scores for the three conditions in the short and long contours.



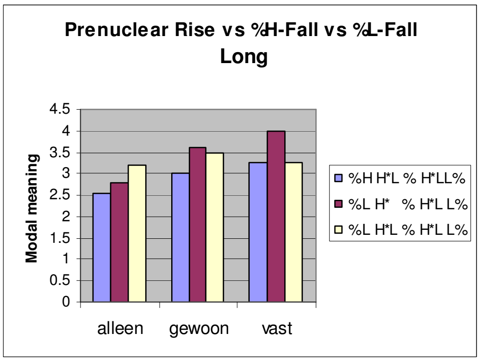


Figure 5. *Mean modal-scores for three contours for three ambiguous words separately in the short and long conditions.*

A repeated measures MANOVA was run to test the null hypothesis that the contour before the boundary does not affect the interpretation of the ambiguous words, with PITCH ACCENT (3 levels), WORD (3 levels) and LENGTH (2 levels) as independent variables. Mauchly's test for spericity was significant for the PITCH ACCENT × WORD interaction, but the interaction itself was not significant. Only main effects were found, for PITCH ACCENT (F [2, 14.51] = 6.35, $p < .01$) and for LENGTH (F [1, 12.84] = 5.06, $p < .05$). Pairwise comparisons revealed that there was no significant difference between the first and third contours, but that the first contour was significantly different from the second and the second was significantly different from the third ($p > .05$). That is, the %L H* % contour in the preboundary phrase attracts significantly more modal interpretations than either %H H*L % or %L H*L %.

3. Conclusions

The generalization that the lexical interpretations of ambiguous Dutch words like *alleen, gewoon* and *vast* is favoured by the presence of a boundary between these words and the next pitch accent in the sentence as well as by the identity condition on the pitch accents on either side of that boundary was supported by a perception experiment in which listeners were asked to choose between a lexical and a modal interpretation of these words. The manipulation of the f_0 of naturally spoken utterances was enough to swing judgements in the direction of a modal meaning if the manipulation involved the removal of the intonation phrase boundary. In one out of nine cases, this shift did not occur. This concerned the word *vast* 'surely' in the interrogative intonation, which finding could be attributed to the incompatibility of the modal meaning and the meaning of the intonation contour. In the case of the pitch accent H*LH, the pre-nuclear 'fall-rise', the results suggest that low pitch after an accentual peak does not imply an intonational boundary if the pitch rises to the next accent, but does if it remains flat until the next accent. This result suggest that H*L as well as H*LH occur as pre-nuclear pitch accents, on the assumption that the last tone is pronounced just before the first tone of the next pitch accent if there is no boundary, but immediately after the preceding tone of the same pitch accent if there is no boundary. Second, the fall of an accentual peak appears to behave like a rise-fall when preceding a rise-fall, while the rise-part does not. This finding shows independently that H*L is indeed a pitch accent of Dutch, and that an accentual peak is to be analyzed as resulting from H*L, with a preceding L-tone coming either from a boundary %L or from the trailing L-tone of a preceding H*L pitch accent. An alternative whereby the peak is analyzed as LH*, with the fall being due to a following L-boundary tone, as offered by MAE ToBI, is thus to be rejected.

Acknowledgments. The author wishes to thank Judith Hanssen for recording the source utterances, Joop Kerkhoff for assistance with the production of the cd-roms containing the experiment, Renske Teeuw for conducting the experiments, and Jörg Peters for help with the statistical processing of the scores.

4. References

[1] Bolinger, D., 1998. *Intonation and its Uses: Melody in Grammar and Discourse.* Stanford: Stanford University Press.

[2] Wells, J.C., 2006. *Intonation: An Introduction.*Cambridge: Cambridge University Press.

[3] Gussenhoven, C., 2005. Transcription of Dutch Intonation. In *Prosodic Typology: The Phonology of Intonation and Phrasing.* Sun-Ah Jun (ed.). Oxford: Oxford University Press, 118-145

[4] Beckman, M. E.; Hirschberg, J.; Shattuck-Hufnagel, S., 2005. The original ToBI system and the evolution of the ToBI framework. In *Prosodic Typology: The Phonology of Intonation and Phrasing.* Sun-Ah Jun (ed.). Oxford: Oxford University Press, 9-54.

[5] Gussenhoven, C., 2004. *The Phonology of Tone and Intonation.* Cambridge: Cambridge University Press.

[6] O'Connor, J.D.; Arnold, G.F., 1973. *Intonation of Colloquial English*. 2nd edition. London: Longman.