

An Experimental Investigation of Meaning Differences between the ‘Early’ and the ‘Late’ Accent-Lending Fall in Dutch

Johanneke Caspers

1. Introduction

The experiments presented in this paper form part of a larger project investigating the meaning of Dutch melodic shapes. We know that intonation plays a role in the marking of pragmatically prominent elements (focus) and in the location of boundaries (see for example Bolinger 1986, 1989, Cutler 1991, Ladd 1996, Pierrehumbert and Hirschberg 1990), but we know little about the functional differences between the different types of melodic boundary markers and different types of accent-lending pitch configurations. The present paper focuses on possible meaning differences between two types of accent-lending falls.

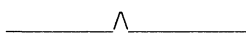
There are two well-known models of Dutch intonation: the GDI or Grammar of Dutch Intonation as developed by 't Hart, Collier and Cohen (1990) and the autosegmental model formulated by Gussenhoven (1984, 1988, 1991). Gussenhoven's model distinguishes — at least — two different types of falling pitch accent (%H H*L and %H !H*L),¹ whereas the GDI recognizes only one type of accent-lending fall ('A'). The aim of the experiments presented below is to find out whether Dutch has one or two types of falling pitch accent.

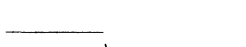
The GDI distinguishes ten so-called perceptually relevant pitch movements: five types of pitch rises ('1'–'5') and five types of pitch falls ('A'–'E'), differing along the dimensions *rate of change* (abrupt or gradual), *size* (full or half) and *timing* with respect to syllable boundaries (early, late or very late). Fall 'A' is described as an abrupt full late fall, lending accent to the syllable it is associated with. It usually appears in either of two very frequent pitch patterns: the 'pointed hat' or the 'flat hat' contour. The pointed hat contour ('1&A') consists of an accent-lending rise ('1') immediately followed by an accent-lending fall ('A'), resulting in one accented syllable, whereas the flat hat contour ('1(0)A') marks two syllables, corresponding to the position of either the rise or the fall, as accented (the peaks are connected by a stretch of level high syllables, '0'). In the implementation of the GDI for Dutch text-to-speech synthesis (Collier 1991), 'A' starts at 20 ms before the vowel onset when it occurs in a flat hat contour, whereas it starts at 80 ms after

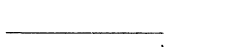
the vowel onset in a pointed hat (cf. table II). This means that the timing of the fall with respect to the accented syllable may depend on the melodic environment it occurs in (in divergence from the standard description of 'A' as given in 't Hart et al. 1990).

In Gussenhoven's autosegmental model, a flat hat contour is described as a sequence of two H*L tone morphemes (each consisting of a high tone, associated with the accented syllable, followed by a low tone), where the prenuclear accent loses its L as a result of a complete linking of the two tone morphemes (H*L H*L → H* H*L). The second H*L tone may be downstepped (H* !H*L), which leads to an earlier timing for the fall than in a 'normal' H* H*L contour (cf. Gussenhoven 1991:147–9, Rietveld and Gussenhoven 1995:376–7). This means that the autosegmental approach to Dutch intonation distinguishes two categorically different types of fall. There is a meaning difference associated with this phonological difference: the downstepped fall supposedly sounds more final than the non-downstepped one (Rietveld and Gussenhoven 1995, see also Swerts, Bouwhuis and Collier 1994, Ladd 1996).

In the experiments reported upon in the present paper, an early and a late fall are contrasted with each other and with the default pitch accent (the most frequently used pitch accent type in Dutch, the 'pointed hat' or H*L). Below the three pitch accent types are illustrated with a schematic contour:

- (1)  (0) 1 & A (0) (= notation GDI)
%L H*L L% (= notation Gussenhoven)
Ik heb Jolanda gezien

- (2)  (0) A (0)
%H !H*L L%
Ik heb Jolanda gezien

- (3)  (0) &A (0)
%H H*L L%
Ik heb Jolanda gezien
I have Jolanda seen
'I have seen Jolanda.'

Throughout the remainder of the text, the label '1&A' will be used for the pointed hat; for the early fall the label 'A' is used, and for the late fall '&A', because this movement is phonetically very similar to the fall in the pointed hat.

2. Experimental findings and theoretical meaning analyses

Experimental investigations of single-accent contours '1&A' and 'A' (Caspers 1997, 1998, 1999; Caspers, van Heuven and van Zwol 1998) revealed that the early accent-lending fall ('(Ø)A(0)' or %H !H*L L%) is not particularly suitable to highlight new or unpredictable information, whereas the pointed hat contour is suitable for focusing unpredictable as well as predictable information. Keijsper's GDI-based theoretical analysis of the meaning of Dutch melodic shapes (1984) offers an explanation for the more limited distribution of fall 'A', since this contour (her 'type III') supposedly indicates that the existence of the focused information was projected before the moment of speaking (which is not the case when the information is unpredictable). Furthermore, Caspers (1998) showed that the early fall 'A' makes the speaker sound more irritated than '1&A', but 'A' does not sound more final than '1&A' (nor less).

Grabe, Gussenhoven, Haan, Marsi and Post (1998) investigated the attitudes associated with a number of Dutch intonation contours, differing in onset height and pitch accent type. Their material included the pointed hat and the late fall (but not the early 'downstepped' fall) and results revealed that '1&A' (%L H*L) sounds polite and friendly while '&A' (%H H*L) does not, and that '&A' sounds detached and irritated whereas '1&A' does not (cf. their figure 3).

The experimental findings provide support for Keijsper's proposal that a final pointed hat marks information as *new* and that the early fall indicates that the existence of the focused information was *projected before the moment of speaking*, a difference in *information status* (see for example Gussenhoven 1984, Pierrehumbert and Hirschberg 1990). In Gussenhoven's approach, on the other hand, '1&A', 'A' and '&A' are all instances of the H*L tone morpheme, which means that they supposedly share the abstract meaning of adding *new* information to the background, but differ in *attitude*. The latter implication is largely supported by the experimental results, yet the former implication is not.

The above leads to the following research questions: Do the early and the late accent-lending falls differ in information status? What are the attitudes associated with both types of fall?

3. Hypotheses

Based on the theoretical analysis formulated by Keijsper (1984), on the experimental findings reported above and on my own intuitions, the following hypothesis was formulated:

- a difference in timing of a pitch fall corresponds to a difference between *new* and *projected* information

The early fall ‘A’ is therefore expected to suit information already projected, while the late fall ‘&A’ — like ‘l&A’ — fits new information. Inspired by Judith Haan (p.c.) and Grabe et al. (1998, p.65) it was further hypothesized that:

- a difference in onset height of a contour corresponds to a difference between *elliptic* and *non-elliptic* information

This means that a contour starting with a high onset (e.g. ‘&A’) is expected to refer back to earlier information and therefore fit ‘elliptic’ information, whereas ‘l&A’ does not.

4. Experimental approach

The hypothesized differences between the three contour types under investigation were captured in three context types, represented in table 1.

Table 1. *Context characteristics*

contexts	‘new’	‘projected’
‘plus ellipsis’	1	–
‘minus ellipsis’	2	3

The opposition between ‘new’ and ‘projected’ information was operationalized as an opposition between an utterance forming the second part of a conjunction and the same utterance in isolation, presenting ‘new’ information (for an example, see below). The opposition between ‘plus’ and ‘minus ellipsis’ was operationalized as the difference between utterances implicitly and explicitly contradicting the previous utterance (see below). The ‘plus ellipsis’ condition is not considered

relevant for the utterances containing 'projected' information — hence the empty cell — because it is inconceivable to have an utterance which is projected by an earlier utterance (viz., the first part of a conjunction), and at the same time elliptic (the preceding utterance is omitted).

The materials (12 contexts, 4 stimulus sentences, 3 contour types) were presented to the subjects as a series of short conversations between two teachers (A and B) working at the same school, discussing the upcoming school party. The utterance by A was always of the same type, and the response by B — with whom the subjects had to identify themselves — always contained the stimulus utterance (in italics), but could differ in the following ways (the numbers of the examples correspond to the numbers in table 1):

A Jolanda heeft nog niemand benaderd voor de opruimploeg.
Jolanda has yet nobody approached for the cleaning.squad
'Jolanda hasn't approached anybody for the cleaning squad yet.'

B (1) *Ze heeft Marina gebeld.*
She has Marina called
'She has called Marina.'

(2) Jawel! *Ze heeft Marina gebeld.*
Yes.she.has! She has Marina called
'Yes she has! She has called Marina.'

(3) Jawel! *Ze heeft Jan-Willem gevraagd en ze heeft Marina gebeld.*
Yes.she.has! She has Jan-Willem asked and she has Marina called
'Yes she has! She has asked Jan-Willem and she has called Marina.'

In example (1) B's utterance supposedly contains new information (at least new to the hearer); also, the utterance can be viewed as elliptic, since an explicit refutation of A's statement is omitted. Contour '&A' is expected to fit this context type best (better than '1&A'), whereas 'A' will supposedly not fit. In context type (2) B's reply also contains new information, but the stimulus utterance is not elliptic anymore, so that the pointed hat is expected to fit best. In context type (3), the information contained in B's utterance can be viewed as projected before the moment of speaking and therefore 'A' is expected to fit here. However, since it looks as if virtually all focused information may be interpreted as new, i.e., as an addition to the background shared between speaker and listener (with the exception of vocatives, cf. Caspers 1997, 1999), both '&A' and '1&A' may fit the 'projected'

contexts also. The other way round, it seems far less likely that an utterance containing new information may carry a pitch accent marking the focused information as ‘already projected’.

4.1 *Design*

The investigation comprised two sorts of test: a pairwise suitability comparison experiment and an absolute rating experiment. To avoid a direct influence of the pitch of the context preceding the target utterance, subjects were presented with only visual representations of the dialogue contexts.

In the pairwise comparison experiment subjects were asked to picture themselves as speaker B in each of the (visually) presented dialogue contexts and to decide which of the two melodic versions of the target utterance — which they could make audible as often as they wished — best fitted the given situation.

In the rating experiment subjects were asked to judge each combination of context and contour type (using the same materials) on the following ten-point scales: acceptability, irritation, detachment and finality. ‘Acceptability’ is used as a complement to the pairwise comparison data; ‘irritation’ and ‘detachment’ are two of the semantic scales used by Grabe et al. (1998), and ‘finality’ is an attitude associated with downstep (cf. Swerts et al. 1994, Rietveld and Gussenhoven 1995, Ladd 1996). The subjects were asked to judge the melody of a specific (audible) target utterance in a specific (visually presented) dialogue context on a ten-point scale, ranging from e.g. ‘sounds not final at all’ (‘this is not the end of what the speaker has to say’) to ‘sounds very final’ (‘the speaker has finished’).

4.2 *Method*

Two Dutch intonologists, a male and a female, realized four different target utterances with each of the three intonation contours.

Thirty-six native Dutch listeners participated in the experiments.² Their ages varied between 21 and 58; no hearing difficulties were reported and they were paid. The stimuli were presented to the listeners via an interactive computer program.³ Since the majority of subjects participated through Internet, there was no strict control over the circumstances under which the experiment was performed (such as ambient noise, type of headphones, type of loudspeaker, etc.). Subjects needed approximately 30 minutes to complete the task.

The order of the two experiments as well as the order of the four scales within the rating experiment was counter-balanced over subjects.

5. Results

5.1 *Pairwise comparison test*

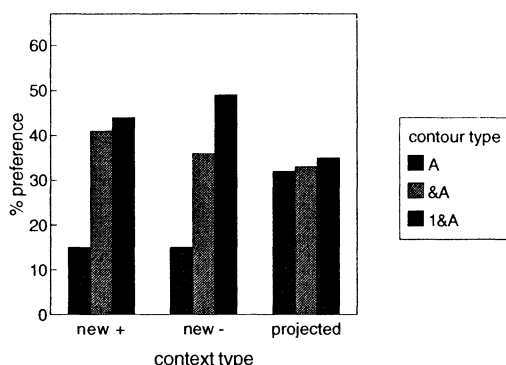


Figure 1. *Percent preference for contour type per context type*

Figure 1 presents the overall preferences for the three contour types, broken down by the three context types ('+' and '-' refer to 'plus' and 'minus ellipsis'). Note that the maximum percentage per contour type is 67% (each of the three contour types is present in only two-thirds of the pairwise comparisons because subjects had to choose between two contours, not between three). In both 'new' context types, the pointed hat scores best, whereas the late accent-lending fall ('&A') is a good second; the early accent-lending fall ('A') is the least popular contour, preferred in only 15% of the 'new' cases. When the focused information is the second part of a conjunction ('projected' information), there is no preference for any contour type ($\chi^2 = .721$, $df = 2$, ins.). There are no significant differences between the 'new plus ellipsis' and 'new minus ellipsis' categories ($\chi^2 = 2.01$, $df = 2$, ins.), but the difference between 'new' and 'projected' information is highly significant ($\chi^2 = 48.586$, $df = 2$, $p < .001$). This means that the strongest context effect lies in the opposition between new and projected information; in the former case, the pointed hat and the late fall are preferred over the early fall, whereas there is no preference for any of the contour types in the latter case. There is no significant effect of the factor 'ellipsis', which means that there does not seem to be an interaction between the height of the onset of the stimulus utterance and whether or not the

utterance can be viewed as elliptic. The results for ‘A’ and ‘1&A’ are in line with the preference judgments as reported in Caspers (1998).

The data indicate that the early accent-lending fall (‘A’) differs from the late fall (‘&A’) and the pointed hat (‘1&A’) in that it does not fit the ‘new’ contexts very well, which can be taken as support for the hypothesis that the early and the late accent-lending fall form different phonological categories (or even tonal morphemes).

5.2 Absolute rating

5.2.1 Acceptability scores

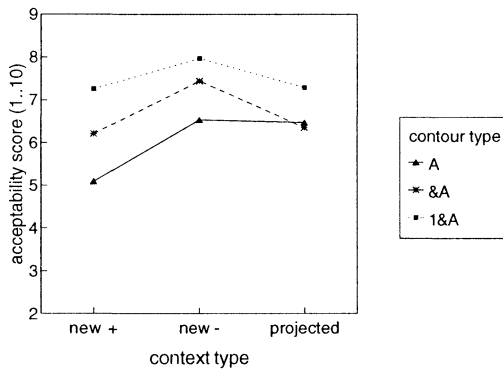


Figure 2. Acceptability scores per contour type, broken down by context type

In figure 2 the acceptability scores are given for the three contour types, broken down by the three context types. There is a significant effect of contour type on the acceptability ratings, $F(2,303)= 11.6$, $p<.001$, and an effect of context type $F(2,303)= 6.7$, $p<.005$, but no significant interaction, $F(4,297)= 1.1$, ins. Post-hoc analyses (Newman-Keuls) reveal that the pointed hat is generally judged as more acceptable than the late accent-lending fall, which in turn sounds more acceptable than the early accent-lending fall. Furthermore, the stimuli in the ‘new minus ellipsis’ contexts are judged as significantly more acceptable than those in the ‘new plus ellipsis’ contexts (which may be attributable to the ‘abrupt’ nature of the elliptic stimuli), but neither differ from the ‘projected’ contexts.

The acceptability ratings confirm the general preference found for the pointed hat in the pairwise comparison test; the fact that the pointed hat and late fall were

preferred over the early fall in the 'new' contexts, but not in the 'projected' contexts (cf. figure 1), is reflected in a relatively higher acceptability of 'A' compared to '&A' and '1&A' in the 'projected' contexts than in the 'new' contexts; however, these differences do not reach significance.

5.2.2 Irritation scores

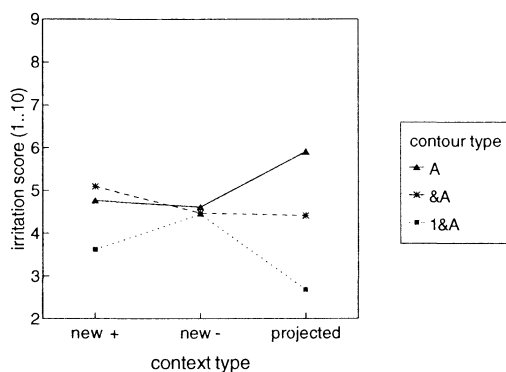


Figure 3. Irritation scores per contour type, broken down by context type

An analysis of variance on the irritation judgments shows an effect of contour type, $F(2,302) = 11.1$, $p < .001$, no main effect of context type, $F(2,302) < 1$, and an interaction between contour and context type, $F(4,296) = 4.2$, $p < .005$. Figure 3 indicates that the early fall ('A') on a 'projected' context sounds more irritated than the late fall or pointed hat, whereas the three contours receive equal scores in the 'new minus ellipsis' context. This impression is supported by post-hoc analyses: when the context contains projected information, 'A' sounds significantly more irritated than '&A', which in turn sounds significantly more irritated than '1&A', while in the 'new minus ellipsis' contexts there is no effect of contour type. Finally, in the 'new plus ellipsis' contexts the late fall sounds significantly more irritated than the pointed hat.

5.2.3 Detachment scores

Figure 4 presents the detachments scores for the various contour types and context types. An analysis of variance shows a main effect of contour type, $F(2,303) = 11.0$, $p < .001$, but no influence of context, $F(2,303) = 1.0$, ins. Post-hoc analyses show that

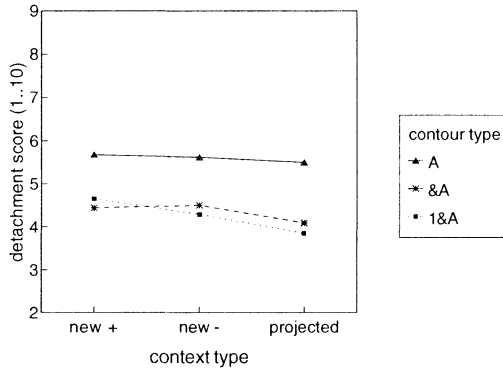


Figure 4. *Detachment scores per contour type, broken down by context type*

‘A’ sounds more detached than ‘1&A’ or ‘&A’. There is no interaction, $F(4,297)<1$. These results indicate that ‘detachment’ is an attitude closely associated with the early accent-lending fall.

Finality scores

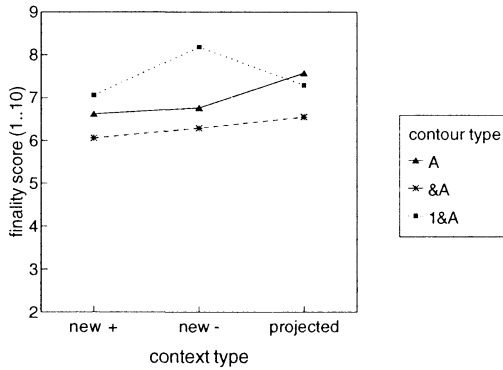


Figure 5. *Finality scores per contour type, broken down by context type.*

The finality scale shows an effect of contour type, $F(2,300)=6.6$, $p<.005$, no difference between the three context types, $F(2,300)=1.7$, ins., and no interaction, $F(4,294)=1.2$, ins. Post-hoc analyses show that the late fall sounds less final than the early fall or the pointed hat.

Leaving the 'new minus ellipsis' condition out of consideration (because the stimulus utterance contains new information but is not initial), the acceptability, irritation and finality scores closely reflect the results reported in Caspers (1998).

6. Conclusion and discussion

The type of experiments described above is rather sensitive to choice of stimulus material and instruction of subjects (cf. for example Caspers et al. 1998, Caspers 1999). Furthermore, one cannot force subjects to perceive stimuli in a specific way, e.g., interpret the second part of a conjunction as 'already projected information' instead of 'new' information (cf. §4). In addition, the various pitch accent types have an important meaning aspect in common (viz., "this is important information"). Bearing these facts in mind, the results of the experiments with regard to the differences between 'A' and '&A' seem quite clear. They can — roughly — be summarized as:

- the early accent-lending fall ('A') does not fit new information, in contrast with the late accent-lending fall ('&A')
- the early accent-lending fall sounds less acceptable, more irritated, more detached and more final than the late accent-lending fall

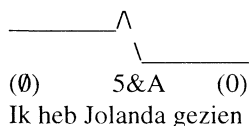
These findings suggest that the two types of fall differ in expressed information status as well as expressed attitude, which may prompt the conclusion that they must belong to two separate linguistic categories.

Obviously, high onset cannot be the reason why the early fall 'A' is unsuited to focus new information, since high onset also occurs with the late fall '&A', which does fit new information quite well. 'A's incompatibility with new information seems to stem from the early alignment of the fall.

The predictions with regard to the differences between the three melodic shapes in terms of information status (cf. §4) were borne out for the 'projected'-'new' opposition — '1&A' and '&A' were preferred over 'A' in the 'new' contexts — but not for the 'plus ellipsis'-'minus ellipsis' opposition — '&A' was not preferred over '1&A' in the 'elliptic' contexts. This may be due to the way the stimulus material was set up: the elliptic version of B's utterance should have been interpreted as a **rejection** of A's statement, which may not have been clear enough, cf. A: "There is nobody from the bar committee yet", B: "*I have seen Jolanda*". However, the tendencies visible in the preference judgments can also be interpreted as weak support for the hypothesis, under a general bias favoring the pointed hat.

Gussenhoven and co-workers seem to be right in the sense that there is a late fall and an early fall and that these shapes differ in attitudes such as ‘detachment’ and ‘finality’; however, a difference in information status is not predicted, whereas the present data indicate that an early fall is not suitable to highlight new information. In my view, the meaning difference between the two types of fall should be regarded as a difference in information status, which leads to secondary (para-linguistic) effects on expressed attitude. For example, the finding that ‘A’ is associated with ‘detachment’ can be predicted from the fact that the early fall cannot readily be used to mark information as ‘new’ (the speaker acts as if the listener already knows — or should know — what the speaker has in mind).

Before attempts are made to supply the GDI with a sixth type of fall, there is an existing melodic configuration to be considered: ‘5&A’, a half rise followed by a full fall (cf. ‘t Hart et al. p. 81):



I think that ‘&A’ and ‘5&A’ are basically the same melodic shape, and further research should provide evidence for this hypothesis. Also, the difference between the pointed hat (‘1&A’) and the late fall (‘&A’), which amounts to a difference in onset height, needs further clarification.

Acknowledgments

This research was funded by the Netherlands Organization for Scientific Research (NWO) under project #300–75–001. Tina Cambier-Langeveld, Vincent van Heuven and an anonymous reviewer are thanked for their valuable comments.

Notes

1. A third type of falling accent within Gussenhoven’s model would be %H L*H, a fall followed by a rise.
2. As a result of a computer crash and a mistake in interpreting the instructions, the data of two subjects could not be used in the analyses.
3. The experiment can be accessed at <http://fonetiek-6.leidenuniv.nl/caspers/le5-intro.html>, programmed by J.J.A. Pacilly.

References

- Bolinger, D.L. (1986) *Intonation and its parts, Melody in spoken English*. Edward Arnold, London.
- Bolinger, D.L. (1989) *Intonation and its uses, Melody in grammar and discourse*. Edward Arnold, London.
- Caspers, J. (1997) 'Testing the Meaning of Four Dutch Pitch Accent Types'. *Proceedings 5th European Conference on Speech Communication and Technology*, Vol. 2, 863–866.
- Caspers, J. (1998) 'Experiments on the Meaning of Two Pitch Accent Types: The 'Pointed Hat' versus the Accent-Lending Fall in Dutch'. *Proceedings International Conference on Spoken Language Processing 1998*, 1291–1294.
- Caspers, J. (1999) 'Experiments on the Meaning of Four Types of Single-Accent Intonation Patterns in Dutch'. [under revision for *Language and Speech*].
- Caspers, J., V.J. van Heuven and N. van Zwol (1998) 'Experiments on the Semantic Contrast between the 'Pointed Hat' Contour and the Accent-Lending Fall in Dutch'. In R. van Bezooijen and R. Kager, eds., *Linguistics in the Netherlands 1998*. John Benjamins, Amsterdam, 65–79.
- Collier, R. (1991) 'Multi-Language Intonation Synthesis'. *Journal of Phonetics* 19, 61–73.
- Cutler, A. (1991) 'Prosody in Situations of Communication: Saliency and Segmentation'. *Proceedings of the XIIth International Congress of Phonetic Sciences*, Aix-en-Provence, Vol. 1, 264–270.
- Grabe, E., C. Gussenhoven, J. Haan, E. Marsi and B. Post (1998) 'Preaccental Pitch and Speaker-Attitude in Dutch'. *Language and Speech* 41, 63–85.
- Gussenhoven, C. (1984) *On the Grammar and Semantics of Sentence Accents*. Foris, Dordrecht.
- Gussenhoven, C. (1988) 'Adequacy in Intonation Analysis: The Case of Dutch'. In H. van der Hulst and N. Smith, eds., *Autosegmental Studies on Pitch Accent*. Foris, Dordrecht, 95–121.
- Gussenhoven, C. (1991) 'Tone Segments in the Intonation of Dutch'. In T.F. Shannon and J.P. Snapper, eds., *The Berkely Conference on Dutch Linguistics 1989*. University Press of America, Lanham MD, 139–155.
- Hart, J., 't, R. Collier and A. Cohen (1990) *A Perceptual Study of Intonation*. Cambridge University Press, Cambridge.
- Keijsper, C.E. (1984) 'Vorm en Betekenis in Nederlandse Toonhoogtecontouren'. *Forum der Letteren* 25, 20–37, 113–126.
- Ladd, D.R. (1996) *Intonational Phonology*. Cambridge University Press, Cambridge.
- Pierrehumbert, J. and J. Hirschberg (1990) 'The Meaning of Intonational Contours in the Interpretation of Discourse'. In P.R. Cohen, J. Morgan and M.E. Pollack, eds., *Intentions in Communication*, 271–311.
- Rietveld, T. and C. Gussenhoven (1995) 'Aligning Pitch Targets in Speech Synthesis: Effects of Syllable Structure'. *Journal of Phonetics* 23, 375–385.
- Swerts, M., D.G. Bouwhuis and R. Collier (1994) 'Melodic Cues to the Perceived "Finality" of Utterances'. *Journal of the Acoustical Society of America* 96, 2064–2075.