

Chapter 5

Transcription of Dutch Intonation

To appear in:

Jun, Sun-Ah (forthcoming) Editor. *Prosodic Typology and Transcription: A Unified Approach*. Oxford University Press.

Carlos Gussenhoven

Centre for Language Studies

University of Nijmegen

September 2000

ABSTRACT

A system for the transcription of Dutch intonation is described, based on the autosegmental description by Gussenhoven. Intonational Phrase (IP) boundaries, Utterance boundaries, IP-final and IP_initial tones, plus pitch accents are described, following the Pierrehumbert model. It uses three initial boundary tones, %L, %H and %HL, three final boundary conditions, L%, H% and % (no tone), five pitch accents H*, L*, H*L, L*H, and H*!H, plus two modifications, Delay (L*-prefix) and downstep (!H). The system has a larger coverage than the IPO grammar, is more transparent than the earlier autosegmental description, while also including the distinction between low rise (L*H H%) and high rise (H*H%), absent from the earlier description. The system is available as an interactive course on the internet at <<http://lands.let.kun.nl/todi>>.

Key words: Dutch, stress shift, accentuation, prosodic hierarchy, initial boundary tone, final boundary tone, optional boundary tone, bitonal boundary tone, downstep, delay, tone modification, pitch accent, high rise, low rise, tonal prefix

Acknowledgements

I am very grateful to Jacques Terken and Toni Rietveld for their competent and supportive collaboration on the design of the ToDI system and the ToDI course. Toni Rietveld deserves special credit for making everybody aware that a system like ToDI was urgently needed, not in the least because of the data collection for the *Corpus Gesproken Nederlands*, a one million word spoken corpus of Dutch, was about to begin. The initial discussions about what ToDI should be like were conducted with the Nijmegen prosody group of 1998, and I thank Esther Grabe, Judith Haan, Joop Kerkhoff, Erwin Marsi, Anneke Neijt, Brechtje Post, and Felix Vieregge for their sensible comments and useful contributions. Valuable input was provided by Aditi Lahiri and her colleagues after a presentation of an earlier version at the University of Konstanz in October 1998. Ludmila Menert and Arthur Dirksen of Fluency Speech Technology (Utrecht) are responsible for the implementation of the first edition of ToDI as an interactive course, which was made available on the internet in July 1999 <lands.let.kun.nl.todi>, while Joop Kerkhoff produced the programs for the second edition (2004). Their work and was supported by a development grant from the Netherlands Foundation for Scientific Research (NWO).

5.1 Introduction

Transcription of Dutch Intonation (ToDI) is a ToBI-like transcription system developed for standard Dutch. A brief introduction to the prosodic structure of this language, which has no lexical tone, can be given on the basis of (1).

- (1) a. Utterance (U)
 |
 Intonational Phrase (IP)
 |
 Phonological Phrase (PP)
- b. Initial boundary tones: %T
 Pitch accents: T*(T)
 Final Boundary tones: T%

The IP corresponds to the ‘tone group’ of the British English tradition, which is closest to the IP as used in ToBI, and is the constituent demarcated by %T and T%. The PP is the domain for clash resolution (to be illustrated below); there is no intermediate constituent between PP and IP. The U is the highest constituent. A minimal instantiation of the elements in (1) would occur on a one-word utterance, like *GefeliciTEERD* ‘Congratulations’, *HALLO* ‘Hello’, or *JA* ‘Yes’, as exemplified in (2), spoken with a perfunctory intonation.

- (2) U{ IP[PP(JA)PP]IP }U ‘Yes?’
 | |
 %L H*

Usually, utterances are more complex. In (3), an example is given of a more elaborate expression, with possible lexical instantiations given in (3a,b,c).

- (3) U{ IP[PP(* *)PP PP(* * *)PP]IP IP[PP(* *)PP]IP }U
 | | | | | | | | | |
 %L H*L H*L H*L H*L H* H% %L H* !H*L L%

- a. De NIEUwe archiTECT / bleef KRAP ZES MAANDEN, maar NIEmand vond dat een proBLEEM
'The new architect stayed less than six months, but nobody minded.'
- b. De TWEEde KEER / kwam er een RAAR, KLEIN MANnetje kijken, in een PAARS PAK
'The second time a strange little man came to watch, in a purple suit.'
- c. Mijn EIgen DOCHter / staat BOven ELke verDENking, zoals u VAST al had verWACHT
'My own daughter is above all suspicion, as you had no doubt expected.'

As these examples suggest, there are no prosodic restrictions on which words appear in which positions. Except for function words with schwa, which are only accented in metalinguistic usage and a few rare idioms, all words are accentable. There are a number of reasons why words are not accented:

- (i) Morphosyntactic reasons. Right-hand constituents of noun and verb compounds are deaccented, for instance. In addition, certain syntactic constituents that are appended at the right edges of clauses typically remain unaccented. Among these are general place and time adverbials, like *hier* 'here' and '*vandaag* 'today', vocatives, and reporting clauses, like *zei Jan* 'said John'. Also, the verb is unaccented when it is adjacent to one of its arguments in an 'eventive' sentence. These conditions may co-occur in the same sentence, as shown in (4). While fully grammatical, this is a somewhat contrived example, as long stretches of unaccented speech are avoided.. *Postbode* is a compound, literally 'post-messenger', *gisteren* a general place adverbial, *heeft een ongeluk gehad* the predicate of *postbode*, *Hendrik* a vocative, and *zei Rie op veelbetekenende toon* a reporting clause.
(4) "De POSTbode heeft gisteren een ongeluk gehad, Hendrik", zei Rie op veelbetekenende toon
the postman has yesterday an accident had Henry said Mary on knowing tone
'The postman had an accident yesterday, Henry, said Mary knowingly
- (ii) Focus structure. The second reason for deaccentuation is that, within the IP, the focus of the sentence is followed by unaccented words, causing the last pitch accent of the intonational phrase to mark the right edge of the focus constituent. This is why the only accent in (5) is on *niets* 'nothing'.
(5) [A: Welke gordijnen passen beter bij de bank?]
['Which curtains would better match the colour of the couch?]
B: NIETS past er bij de kleur van de bank!
'NOTHING matches the colour of the couch!'
- (iii) Phonological. Like French and English, the phonological phrase governs the distribution of pitch accents through rhythmic considerations. Rhythm-induced deletions and additions of pitch accents are particularly relevant for adjective compounds. For instance, *donkerbruin* 'dark brown' has a pitch accent on the second member *bruin* when this is the last pitch accent in its

phonological phrase, as in (6a). The context for this expression must be a conversation about *meubelen* ‘furniture’ to explain the lack of accent on this word. Some adjective compounds tend to have pitch accents on both constituents in this type of position, in particular when occurring in IP-final position, like *ijskoud* ‘ice-cold’, *wonderschoon* ‘wondrously beautiful’, *bloedmooi* ‘stunningly beautiful’, as shown in (6b). When another pitch accented word follows in the same PP, as in (6c,d), only the first constituent is accented, a pronunciation known as ‘stress shift’. The phenomenon which is not restricted to compound words (cf. Visch 1999 and references therein; instead of medial PP-boundaries, there may additionally be IP-boundaries in (6a) and (6c)).

(6) a. $\text{U}\{_{\text{IP}}[\text{PP}(\text{RechtHOEkige meubelen})_{\text{PP}} \text{PP}(\text{kunnen we niet LEveren})_{\text{PP}}]_{\text{IP}} \}_{\text{U}}$
 ‘RECTANGULAR pieces of furniture we cannot SUPPLY’

b. $\text{U}\{_{\text{IP}}[\text{PP}(\text{Ze is BLOEDMOOI})_{\text{PP}}]_{\text{IP}} \}_{\text{U}}$
 ‘She is STUNNINGLY BEAUTIFUL’

c. $\text{U}\{ \text{IP}[\text{PP}(\text{Een RECHThoekige Tafel})_{\text{PP}} \text{PP}(\text{stond bij het RAAM})_{\text{PP}}]_{\text{IP}} \}_{\text{U}}$
 ‘A RECTANGULAR TABLE stood by the WINDOW’

d. $\text{U}\{ \text{IP}[\text{PP}(\text{Een BLOEDmooi MEISje})_{\text{PP}}]_{\text{IP}} \}_{\text{U}}$
 ‘A STUNNINGLY BEAUTIFUL GIRL’

The prosodic hierarchy of Dutch continues downward with the prosodic word, the foot, and the syllable (Booij 1995:143). Not all of these are relevant to intonational structure. The foot, a trochee, is relevant because its head serves as the association site for the pitch accents of the language. Outside ‘stress shift’ contexts (cf. (6)), one foot in a phonological word will attract a pitch accent, and its head is the word stress (cf. Hayes 1995:18ff, Gussenhoven and Bruce 1999). Recent treatments of Dutch word stress are Zonneveld et al. (1999:492-515), Gussenhoven (2002).

5.1.1 Melodic aspects.

The melodic aspects of Dutch intonation are quite complex. Although the number of pitch accents in a PP will rarely exceed three, there is no principled limit, as there is in Bengali, which allows only one (Hayes and Lahiri 1991), and thus no limit to the number of pitch accents in the IP. Moreover, the language offers a large number of pitch accents to choose from, also in prenuclear position, which express various discursual meanings. In this latter respect, it is different from French, which has a choice of only two pitch accents in prenuclear position, or Bengali, which must choose $L^* H_p$ (Hayes and Lahiri 1991, Post 2000). There are constraints on the number of *different* pitch accents within the IP: if there are more than two, all the prenuclear ones are usually the same.

The primary purpose of ToDI is to make a transcription system available for characterising the intonation of Dutch utterances or example sentences, including large amounts of spontaneous speech, from which subsequently generalisations about accentuation, phrasing, and tone choice could be extracted. It represents an improvement over the notation proposed in the IPO grammar (Collier and ʔ Hart 1981, ʔ Hart, Collier and Cohen 1990, ʔ

Hart 1998) in that many more contrasts are catered for. It is more transparent than the description of Gussenhoven (1988, 1991). The same advantage is held over Beckman and Pierrehumbert's (1986) description of American English, from which the ToBI system was derived (Beckman and Ayers 1994). For instance, ToDI uses no abstract tones of the sort L-H% to mean 'mid level pitch', but has only tones for which individual phonetic targets can be identified.

Before presenting the system, a number of features are highlighted in section 2 that distinguish it from ToBI. Section 3 then presents and exemplifies the main nuclear and pre-nuclear contours. In section 4, a brief account is given of contours that are predicted by ToDI, if it is viewed as an orthogonal system, but of which we have no recorded examples. Finally, section 5 systematically compares ToDI notation with the notation used in the IPO grammar ('t Hart, Collier and Cohen (1990), my own earlier description, and American English ToBI.

5.2 Some less commonly adopted conventions

ToDI has a number of features that may be unexpected for users of other ToBI-like systems. First, the system only covers the 'To' part of ToBI: prosodic breaks are only included from the tonally marked constituent (the IP) onwards. Of course, some version of the ToBI break indices can always be combined with the ToDI transcription. BI3, ToBI's IP-boundary, can be used for the PP-boundary, and BI1 and BI2 can be used as in ToBI. Second, the final boundary tones of the IP, the only tonally marked constituent in the system, are optional. With two tones, this leads to a three-way opposition in the way IP's end, as opposed to a four-way opposition when a two-layered structure with obligatory tones is assumed. Third, some extra-sentential constituents form accentless IP's. Fourth, there are no leading tones in pitch accents. Fifth, the tones of bitonal pitch accents need not be realised close together in time: prenuclear bitonal pitch accents may define quite long gradual slopes, depending on the distance to the next pitch accent. And lastly, singleton pitch accent H* and L* describe high level and low level pitch, respectively, rather than a single high and low target from which to rise or fall. Below, we discuss each point in turn.

5.2.1 A single tonally marked phrase

The intonational phrase is the only prosodic constituent to be marked by boundary tones.¹ Its beginning is always marked by a boundary tone, like L% (see section 2.2). At the end, the boundary tones L% and H% may appear. Example (7) illustrates the occurrence of %L and L% with the pitch accent H*L, the most neutral contour of the language (cf. 't Hart et al. 1990).

(7)



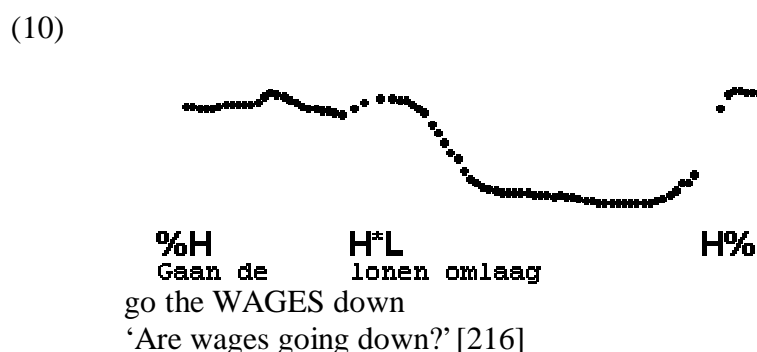
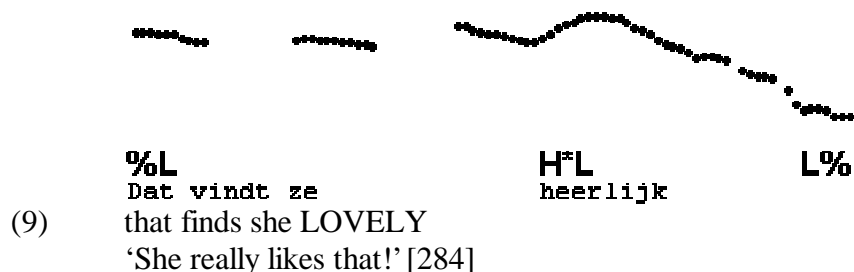
TAKE then so the time
 'So then take the time' [074]

In having only a single tonally marked constituent, ToDI differs from the two-phrase intonational structure proposed by Beckman and Pierrehumbert (1986) for English, with an intermediate phrase and an intonational phrase both contributing final boundary tones. In van den Berg, Gussenhoven and Rietveld (1992), Gussenhoven and Rietveld (1992) it was assumed that, in addition to the IP, the Utterance was provided with boundary tones in Dutch. Specifically, any occurrence of H*L L% and L*H H% was taken to mark the end of an Utterance, while H*L H% was taken to either mark the IP-end. The effect was that only Utterance-final IP's could end in extra low pitch (the effect of LL%) or extra high pitch (the effect of HH%). This assumption was made in the interest of a synthesis-by-rule system for intonation, NIROS, which simulates intonation contours for read speech and which necessarily represents an idealisation of speaker behaviour. However, in spontaneous speech, there is no indication that the regularity holds. For German, which closely related to Dutch, Grabe (1998:171) found HH% as readily in final as in non-final IP's.

5.2.2 Boundary tones

ToDI recognises three ways in which unaccented syllables can be pronounced at the beginning of the IP. One of these is mid or low pitched, and clearly represents the neutral option; a second is high pitched, a marked option, though it is not uncommon before low pitched accents. The third is a rare, highly marked falling pattern. These patterns are transcribed as in (8), and illustrated in (9), (10) and (11).

- (8) Initial boundary tones: %L
%H
%HL



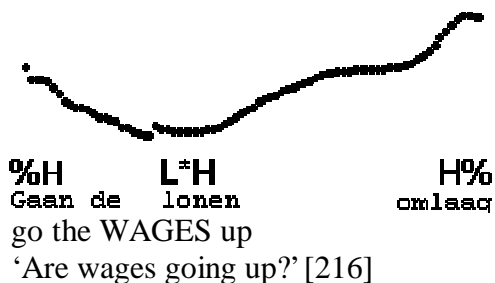
(11)



we had no IDEA of how it FURTHER must-PAST
'We had no idea how to go on' [234]

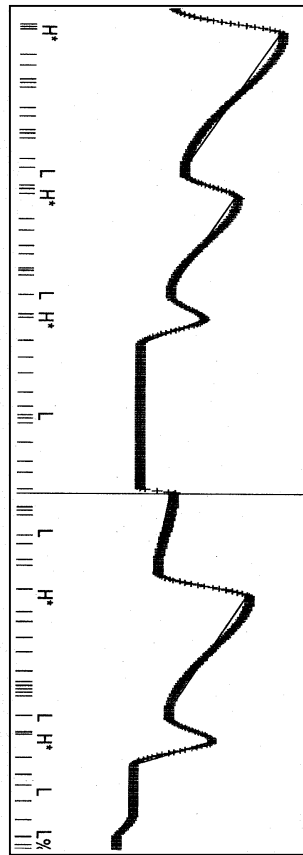
There are two final boundary tones, H%, illustrated in (10) and (12), and L%, illustrated in (7), (8) and (11) above.

(12)



'Are wages going up?' [216]

Final boundary tones are optional leading to three phonological right edges of the IP. ToDI's precursor, Gussenhoven (1988), took these boundary tones to be matter of automatic spell-out, and the nuclear contours in (7) and (12) were simply transcribed H*L and L*H, respectively. It was only when this description was implemented in NIROS (cf. Gussenhoven and Rietveld 1992) that these effects came to be represented by means of separate tones, and the absence of these tones came to be the representation of 'half-completion', which would occur in the mid-ending counterparts of (7) and (11), illustrated in (16) and (24) below. An example of the contrastive absence of a boundary tone from Gussenhoven and Rietveld (1992) is reproduced in Figure 1. The advantages of making T% optional were convincingly argued for by Grabe (1998) for German and English, who introduced the notation 0% to signal absence of tone at an IP-boundary.



*Figure 5.1. Example of a rule-based synthesised contour on PTT Telecom Inlichtingen Telefoonnummers Binnenland ‘PTT Telecom Information Telephone-numbers Interior’, consisting of five H*L pitch accents in two IPs with downstepping accents, one with H*L in Utterance non-final position and one with H*L L% in Utterance-final position. From Gussenhoven and Rietveld (1992).*

accents, a view which has proved useful in descriptions of German and English (Gussenhoven 1983, Féry 1993, cf. ‘displacement’, which term may be preferable to ‘linking’, as the latter term is also in use in the sense of ‘association’, Grabe 1998). Figure 2 illustrates the concept graphically for three pitch accents: the first column gives nuclear realisations, the second prenuclear ones (after Gussenhoven 1983).

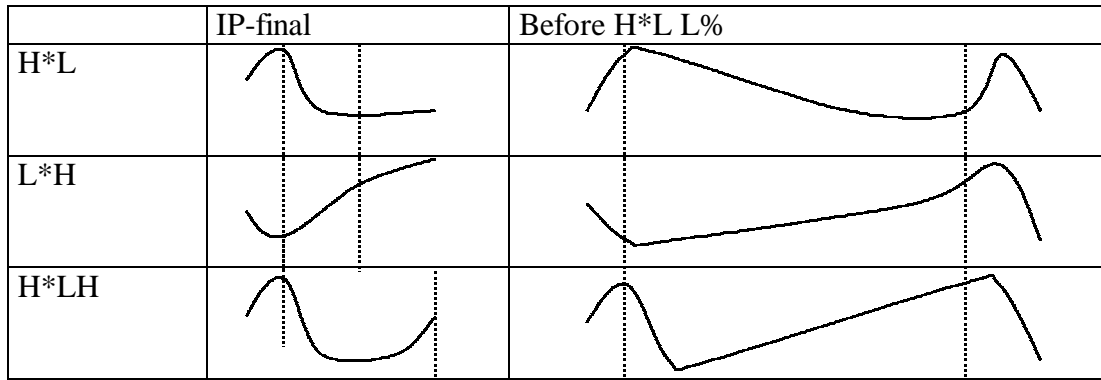


Figure 5.2. Schematic illustration of the effect of (partial) Tone Linking. The dotted lines mark off the contour sections defined by the three pitch accents in final and nonfinal position in the IP (after Gussenhoven 1983).

5.2.6 Singleton T* describing level pitch

ToDI's singleton H* and L* describe accented high and low pitch targets, as in other systems, which, unlike what is intended in other systems, continue that pitch target until a new tone is transcribed., i.e. H* and L* ‘spread’. Falling and rising pitch are transcribed with H*L and L*H, respectively.

5.3 ToDI: the contours

ToDI was designed to handle all the contrasts presented in ' t Hart, Collier and Cohen (1990), Gussenhoven (1988), plus the contrast between ‘high rise’ and ‘low rise’ (Gussenhoven and Rietveld 2000). Notation devices for additional features, such as the difference between a normal and a raised peak, can of course be added to the system at the discretion of the researcher. An overview is given in (14).

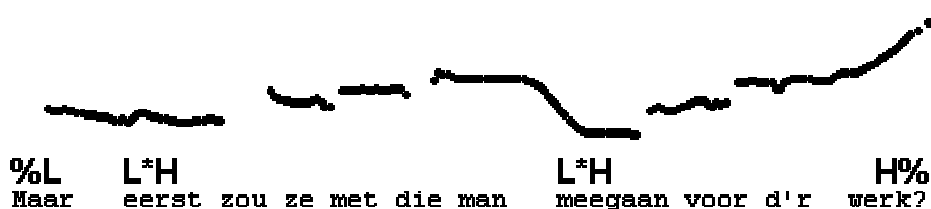
- (14) Initial boundary tones: %L
(repeated from (8)) %H
%HL
- Final boundary tones: L%
(Optional) H%
- Pitch accents: H*
L*
H*L
L*H
H*!H



- (17) that does-it with his ELEPHANT'S -TRUNK of-course
 'It does that with its trunk, of course' [029]

The low rise

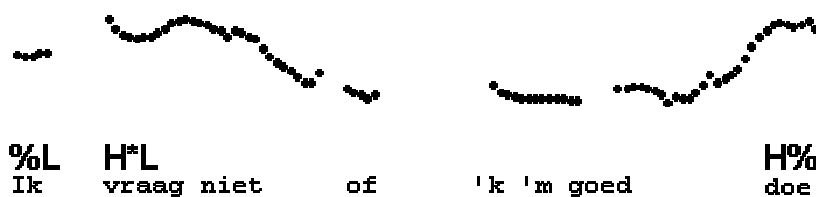
Example (18) has two L*H pitch accents, followed by H%. Again, the H-tone of the prefinal pitch accent is realised just before the L* of the next pitch accent, in this case causing a gradual rise, whose slope will vary without obvious implications for the identity of the contour. The part described by L*H H% is termed the 'low rise'.



- (18) but FIRST would she with that man ACCOMPANY for her work
 'But was she first planning to join that man on a business trip?' [364]

The fall-rise

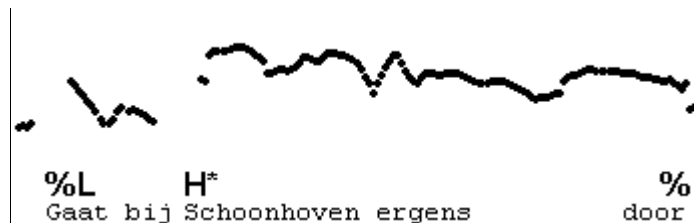
The fall-rise, H*L followed by H%, is a quite frequent contour of Dutch, both in Utterance-final and nonfinal position, as illustrated in (19) and (20), respectively.



- (19) i ASK not whether i it good do
 'I'm not asking if I'm d oing it right' [163]

The level contour

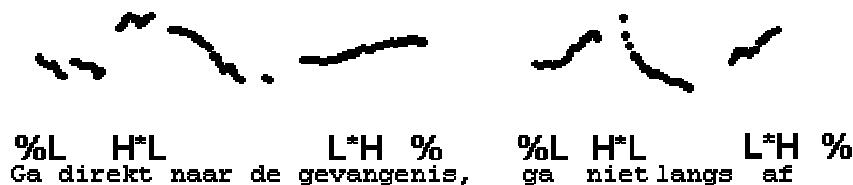
The most prototypical ‘listing’ intonation is one that has mid pitch in the accented syllable, which continues at more or less the same level until the IP-end. It is transcribed H* %, and shown in (24).



- (24) goes near Schoonhoven somewhere through
 ‘... takes a shortcut near Schoonhoven or thereabouts, ...’ [01 8]

The half-completed rise

Like the low rise, the half-completed rise has low pitch in the accented syllable, then rises immediately, but unlike the low rise does not rise further in the final syllable. It is transcribed L*H %, as in (25), where it occurs twice in nonfinal IP’s. It is readily usable as a ‘listing intonation’, as an alternative to the level contour. Phonologically, the difference between L*H H% and H* H% parallels that between L*H % and H* %.

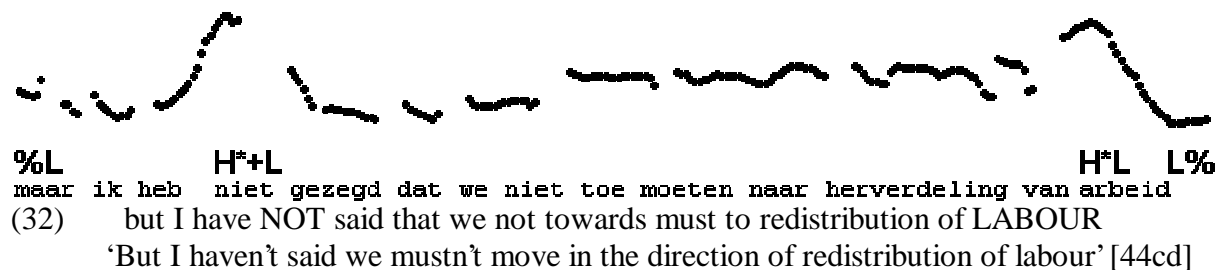


- (25) go DIRECTLY to the JAIL go NOT along OFF
 ‘Go directly to jail, don’t pass *Go*, ...’ [302b]

Vocative chant

The vocative chant is a contour with minimally two pitch levels, frequently accompanied by lengthening of the initial syllables of each level (Gussenhoven (1993). It is transcribed H*!H %, as in (26), where it is used in a nonfinal IP. In multi-level realisations, a contour-type that does not exist in English and forms a cascading sequence of non-accent-lending pitch levels, we simply transcribe H*H!, as in (27), where five levels are formed, on *moet niet me-*, *teen het*, *antwoord*, *ge-*, and *-ven*, respectively.

convention that the last tone of a polytonal pitch accents moves off to the right, where it creates a target just before the next pitch accent (see 5.2.5). The pitch accent H*+L only appears contrastively in prenuclear position, where it is typically followed by a gradual rise to H*. For example, in (32), we have H*+L on *niet* ‘not’, followed by high pitch on *arbeid*. This contour is clearly distinct from one in which the high pitch on *niet* is followed by a gradual fall, as illustrated in (33), spoken by the author.



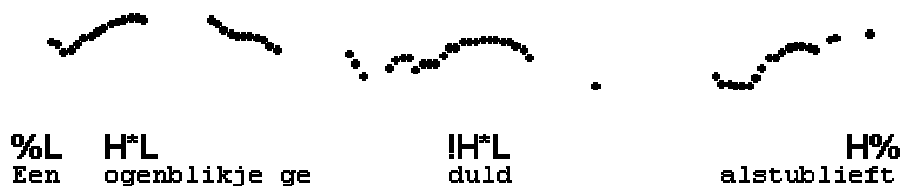
(33) [274]



The explanation of the restrictive distribution of H*+L is that this pitch accent is really the occurrence of the H*LH pitch accent (see Gussenhoven 1983, 1988, and Figure 1). This pitch accent resolves as H*L H%, the fall-rise, in IP-final position, but in prenuclear position it is subject to Tone Linking, i.e. undergoes a rightward shift of the last tone, H, as expected. The L-tone, which is nonfinal within the pitch accent, does not shift, and defines the low point of the steep fall from H*. This not only explains why H*+L is followed by a gradual rise, but also why contours like H*+L H*L are readily replaceable with H*L H% %L H*L (cf. Cruttenden 1994: 59), from which they must be historically derived. In (33), the internal IP-break would occur after *gezegd*, the last word of the main clause, which is not the location of the target of L, suggesting L is not some kind of boundary tone. So H*+L is just a more surface-true notation of prefinal H*LH. (In the second edition, the symbol H*LH will be substituted for H*+L.)

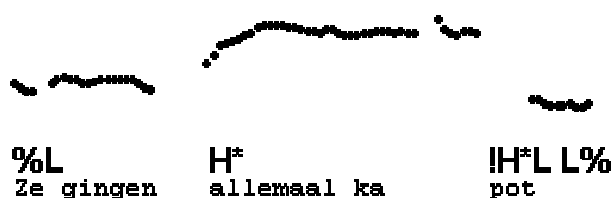
5.4.3 Downstepped contours

In (26) and (27), there occurred an example of a downstepped tone, the unaccented trailing tone of the vocative chant H*!H %. As shown in (27), each of the levels is lower in pitch than the one before. The more usual tone to be downstepped is H*. Two types of contour are particularly frequent. In the first, the downstepped H* occurs after H*L, as in (35): the pitch sags between the earlier high peak and the later low peak, which is transcribed by the trailing L-tone of the first pitch accent.



- (35) a MOMENT-DIM PATIENCE please
 'Just a moment, please'

The second type is probably more frequent, and was termed the 'flat hat' by the IPO intonologists (t Hart, Collier and Cohen 1990). The first accent in (36) has H*, which maintains its pitch until the downstepped !H* of the second accent, causing an early fall which starts just before the final accented syllable. The pitch will variably fall from mid to low or may be low throughout, without there being a noticeable difference. One of the most frequent transcription errors we have encountered is a failure to mark final accents in contours like (35) as accented. In such cases, transcribers incorrectly interpret the contour as one with a single (contrastive) accent on the word which has the first accent.



- (36) they went ALL BROKEN
 'They were all broken [122]

Examples (37) and (38) show these contours with multiple accents, with cascading series of pitch targets for the downstepping !H*-tones. Example (38) replicates an example in Collier and t Hart (1981).

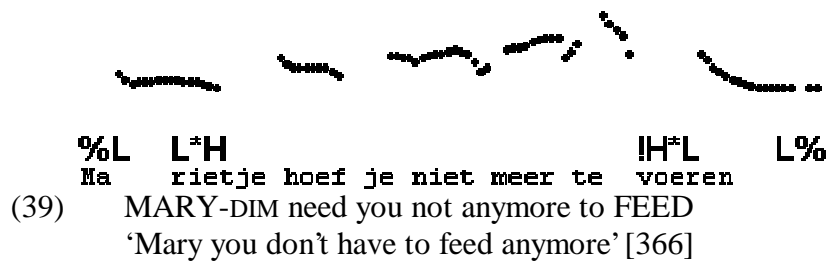


- (37) ALL those COMPLICATED RULES have-been ABOLISHED
 'All those complicated rules have been abolished' [128]

- (38) [321]



Downstepped !H* may occur after any pitch accent containing a H-tone. In (39), there is one after L*H.



Downstepped !H*, lastly, may occur after %H. That is, there are one-accent contours with downstep. This type of contour is not improbable in utterances that represent titles of stories read aloud. Example (40) might be such an utterance.



5.4.4 Delayed accents

In Gussenhoven (1983), I proposed a modification [DELAY] which can affect the three basic pitch accents of that description, H*L, L*H and H*LH. The delayed version of L*H is a contour that ToDI represents as L* H%, the low rise (cf. 5.3.1).² The delayed H*L, which creates a low-pitched accented syllable followed by a rise-fall, either inside that syllable if it is IP-final, or in the next syllable if it is not, is transcribed L*HL in ToDI. The extraneous status of the prefixed L* is nicely illustrated by the fact that the original starred tone H* does not get copied into the accentless IP's discussed in 2.3. That is, after L*HL H%, for instance, the reported clause has L-H%, not HL-H%.

Examples (41) and (42) illustrate this pitch accent, in combination with different boundary conditions.

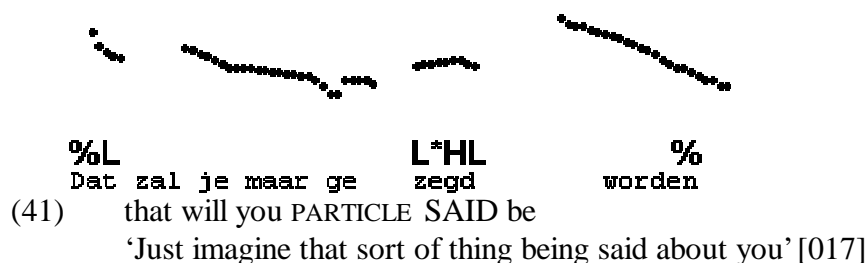


Table 1. Full set of nuclear pitch accents in three boundary conditions generated in the ToDI system, plus prose labels. No examples are available for the italic labels, but it is believed that these represent well-formed contours of Dutch.

| | L% | H% | % |
|-------|-----------------------------------|-------------------------------|---|
| H*L | Fall | Fall-rise | Half-completed fall |
| !H*L | Downstepped fall | Downstepped fall-rise | Downstepped half-completed fall |
| H* | <i>High plateau- Slump</i> | High rise | Level |
| H*!H | <i>Low-ending vocative chant</i> | <i>Vocative fall-rise</i> | Vocative chant |
| L*H | <i>Delayed high plateau-slump</i> | Low rise | Half-completed rise |
| L* | <i>Falling low level</i> | Low low rise | <i>Low level</i> |
| L*HL | Delayed fall | Delayed fall-rise | Delayed half-completed fall |
| L*!HL | Downstepped delayed fall | Downstepped delayed fall-rise | Downstepped delayed half-completed fall |

First, the (delayed or non-delayed) high plateau-slump (the term is from Cruttenden 1994) is a contour that starts high in - or after, in the delayed version - the accented syllable and continues high until a fall to low just before the final syllable. Second, the (falling) low level contours would seem to be used when repeating someone else's statement in a scathing manner. Third, the low-ending vocative chant is a version of the vocative chant that expresses a high degree of impatience, while the vocative fall-rise is just like the vocative chant, with a final rise on the final (lengthened) syllable (the latter was given for English in Gussenhoven 1983, as example (36)).

5.6 Comparison with other systems

It would be appropriate at this point to discuss the way other descriptions have dealt with the contours presented in the above sections. Such a discussion not only define the increase in descriptive coverage, but may also be helpful for those who would like to re-transcribe into the ToDI system contours in the literature, and facilitate the identification of ToDI contours by those who are more familiar with other conventions. However, the space a full treatment would require is prohibitive in the context of this chapter. It is possible, though, to give these comparisons in tabular form, and dispense with a discussion of the advantages and disadvantages of the various systems. This is done in this section.

The first comparison is with the IPO grammar, the second with the autosegmental description (Gussenhoven 1988, 1991, Van den Berg et al. 1992, Gussenhoven and Rietveld 1992), and the third with ToBI, as proposed for American English (Beckman and Ayers 1994), whose intonational system is very similar to that of Dutch.

5.6.1 The IPO grammar compared with ToDI

Table 2 lists ToDI transcriptions for nuclear pitch accents plus boundary tones that were illustrated in the sections above, together with the IPO transcriptions used to transcribe those contours.³ It is clear that the IPO grammar has fewer contours, and many ToDI contrasts are simply not expressed (Gussenhoven 1988). As stressed in that publication, however, the IPO grammar was never designed to be an exhaustive description. It was based on a corpus of speech, and there is of course no guarantee that in any given corpus all possible contours are actually attested. The IPO researchers also put a lower limit of 6% on the frequency of occurrence of any one contour for it to be included in their grammar.

In Table 2, some gaps in the IPO column have been filled with transcriptions that Hart, Collier and Cohen (1999) do not give, but that could be used to represent the contours concerned. These non-standard transcriptions are given in parentheses.

Table 2. Nuclear contours in ToDI compared with their counterparts in IPO, with prose labels as used in this chapter. Bracketed IPO transcriptions are not produced by the IPO grammar, but might be if it were adapted.

| ToDI | IPO | Prose label |
|----------|---------|--|
| H*L L% | A | Fall ^a |
| H*L % | (1 C) | Half-completed fall |
| H*L H% | A 2 | Fall-Rise |
| !H*L L% | A | Downstepped fall ^a |
| !H*L % | - | Downstepped half-completed fall |
| !H*L H% | - | Downstepped rise-fall |
| H* % | 1 | Level tone |
| H*!H % | 1 E | Vocative chant |
| H* H% | 1 2 | High rise ^b |
| L*H H% | (3 2) | Low rise |
| L*H % | (3) | Half-completed rise |
| L* H% | (2) | Low low rise ^c |
| L*HL L% | (3&B) | Delayed fall |
| L*HL % | 3 C | Half-completed delayed fall |
| L*HL H% | (3&B 2) | Delayed rise-fall |
| L*!HL L% | - | Delayed downstepped fall |
| L*!HL % | - | Delayed downstepped half-completed fall |
| L*!HL H% | - | Delayed downstepped fall-rise ^d |

- IPO 'A' corresponds with H*L L% in all contexts except in the 'flat hat' contour, which is transcribed '1 A'. This contour is equivalent to H* !H*L L%, i.e. with downstepped H*, and hence phonetically with an earlier fall. In t Hart, Collier and Cohen (1990), 'B' is often used for the late variant of 'A'. The 'pointed hat' is '1&A', and is equivalent to ToDI %L H*L L%.
- IPO '1 2' is H* H% rather than L*H H%, both in the description and in the examples given in Collier and t Hart (1981), while IPO '1' is L*H % in Collier and t Hart (1981). H* % and L*H % could be distinguished in IPO as '1' and '3', if '3' were used for H* %.
- IPO '2' is described as a non-accent lending final rise. The grammar in t Hart, Collier and Cohen (1990) generates a singleton '2', however, which could be used to describe this contour.
- None of the contours with italic labels in Table 1 could readily be accommodated in IPO notation. They were omitted from Table 2.

The notations for the prenuclear contours are compared in a similar fashion in Table 3.

Table 3. Pre-nuclear contours in ToDI compared with their counterparts in IPO. Bracketed IPO transcriptions are not produced by the IPO grammar, but might be.

| ToDI | IPO | Prose label |
|-------------|----------------------|-------------------------------------|
| H* | 1 | Prenuclear high |
| !H* | E | Downstepped prenuclear high |
| H*L | 1D | Prenuclear fall |
| !H*L | - | Downstepped prenuclear fall |
| L* | 0 ^a | Prenuclear low |
| L*H | 4 | Prenuclear rise |
| L*!HL | - | Downstepped delayed prenuclear fall |
| H*+L | 1&A 4 5 ^b | Prenuclear fall rise |

- a. IPO cannot mark low-pitched accents in low-pitched surroundings, since it is based on movements.
- b. The ‘5’ represents a (non-obligatory) small extra rise above the usual level just before the next accent, which could be interpreted as the movement from the target of H to that of H*, if H*+L is replaced with H*LH of Gussenhoven (1988). That is, IPO ‘1&A 4’ would equally be transcribed H*+L. As Jacques Terken points out (personal communication), IPO could also use ‘5’ to describe the heightened peak in (30), a feature for which ToDI has no transcription.

5.6.2. The earlier autosegmental description of Dutch compared with ToDI

ToDI bears a strong resemblance to my earlier description. The only contrast recognised in that description which cannot be reproduced in ToDI is that between a plain fall-rise and a half-completed fall-rise. The half-completed fall-rise has mid pitch, instead of low pitch, between the high pitch of H* and the high pitch of H%, something that gives the utterance a tentative ring. This contrast was available in Pierrehumbert (1980) as H* L-H% (plain) vs H*+L H-H% (half-completed), but no provision for it exists in ToBI. It is indeed arguable whether it should not be explained as phonetic variation.

The speech synthesis programme NIROS was designed to take tones and modifications as input in a rather abstract form, and produce a surface phonological representation after the application of automatic spell-out rules (Gussenhoven and Rietveld 1992). As shown in Table 4, ToDI is very close to these latter representations. The modifications were features on the IP, and had effect on the some or all of the pitch accents in it. Thus, > delays all accents, = produces a half-completed version of the last pitch accent, ! causes all H*'s except the first to be downstepped, and &, 'narration', causes all T*'s to spread right, pushing the trailing tone up against the T* or the IP-boundary. (No provision was made for 'stylization', which requires additional durational manipulation.) For instance, input representation (45) would be transformed into (46), which latter representation feeds the phonetic implementation. The contour has a gradual slope from the first accent (pre-nuclear rise), followed by an early fall on the second (downstepped fall).

(45) {!(L*H H*L)}

(46) {%L L* H !H*L L%}

The main addition to the earlier description is H* H%, i.e. the contrast between the low rise and the high rise. Among the pre-nuclear pitch accents, H*+L corresponds to H*LH, as explained in 4.2.2. Of the contours with italic labels in Table 2, the earlier description has no way of creating the low-ending vocative chant. Neither could it produce low level contours. It described the high-plateau slump by means of the narrated H*L, and the vocative chant as the stylised fall-rise. Neither of these was implemented.

Table 4. Nuclear contours in ToDI compared with their counterparts in the earlier autosegmental description with prose labels as used in this chapter.

| ToDI | Autosegmental | Spell out | |
|----------|---------------|-----------------|---|
| H*L L% | (...H*L) | H*L L% | Fall |
| H*L % | =(...H*L) | H*L % | Half-completed fall |
| H*L H% | (...H*L) | H*L H% | Fall-Rise |
| !H*L L% | !(.. H*L) | !H*L L% | Downstepped fall |
| !H*L % | =!(...H*L) | !H*L % | Downstepped half-completed fall |
| !H*L H% | !(...H*LH) | !H*L H% | Downstepped rise-fall |
| H* % | stylised L*H | <i>no impl.</i> | Level tone |
| H*!H % | stylised H*L | <i>no impl.</i> | Vocative chant |
| H* H% | - | - | High rise ^a |
| L*H H% | L*H | L*H H% | Low rise |
| L*H % | =(..L*H) | L*H % | Half-completed rise |
| L* H% | &(...L*H) | L* H% | Low low rise ^b |
| L*HL L% | >(...H*L) | L*HL L% | Delayed fall |
| L*HL % | >= (...HL*) | L*HL % | Half-completed delayed fall |
| L*HL H% | >(...H*LH) | L*HL H% | Delayed rise-fall |
| L*!HL L% | >!(...H*L) | L*!HL L% | Delayed downstepped fall |
| L*!HL % | >=!(...HL*) | L*!HL % | Delayed downstepped half-completed fall |
| L*!HL H% | >!(...H*LH) | L*!HL H% | Delayed downstepped fall-rise |

- a. The high rise was incorrectly considered a variant of the low rise (cf. Gussenhoven and Rietveld 1997).
- b. The delayed rise can undergo more or less right-shifting of the rising movement (Gussenhoven 1983). In its most extreme form, it is equivalent to the 'narrated' L*H, which was described as resulting from spreading L* to the IP-end, where it left some space for H. Delay was implemented in NIROS by inserting an extra L-tone before the L* (or H*, in the case of H*L) causing a right-shifting of some 100 ms.

5.6.3 American English ToBI compared with ToDI

The intonation systems of English and Dutch are extremely similar. American English ToBI might therefore have been used for the transcription of Dutch, as suggested by the correspondences in Table 5. There are two reasons for nevertheless proposing a new system. One is that Dutch has contours which American English lacks (and for which thus no transcription is available in ToBI), like the high-plateau slumps (Table 1), or for which ToBI has proposed no transcription, like the low-ending vocative chant. The second is that employing two phrase-types with obligatory right-hand boundary tones has a number of drawbacks, such as creating undesirable intermediate phrase breaks (cf. Ladd 1996: 96 ff., Nolan and Grabe 1997), which ToDI avoids.

Table 5. Nuclear contours in ToDI compared with their counterparts in ToBI.

| ToDI | ToBI | Prose label |
|----------|------------|---|
| H*L L% | H* L-L% | Fall |
| H*L % | H* H-L% | Half-completed fall |
| H*L H% | H* L-H% | Fall-Rise |
| !H*L L% | !H* L-L% | Downstepped fall ^a |
| | H+!H* L-L% | |
| !H*L % | !H* H-L% | Downstepped half-completed fall |
| !H*L H% | !H* L-H% | Downstepped rise-fall |
| H* % | H* H-L% | Level tone |
| H*!H % | H* !H-L% | Vocative chant |
| H* H% | H* H-H% | High rise |
| L*H H% | L* H-H* | Low rise |
| L*H % | L* H-L% | Half-completed low rise |
| L* H% | L* L-H% | Low low rise |
| L*HL L% | L*+H L-L% | Delayed fall |
| L*HL % | L*+H H-L% | Half-completed delayed fall |
| L*HL H% | L*+H L-H% | Delayed rise-fall |
| L*!HL L% | L*+!H L-L% | Delayed downstepped fall |
| L*!HL % | L*+!H H-L% | Delayed downstepped half-completed fall |
| L*!HL H% | L*+!H L-H% | Delayed downstepped fall-rise |

- a. ToBI has two transcriptions for the downstepped fall. The leading H is used to describe a high plateau that runs from a preceding H-tone to the syllable before the downstepped accent. This implies that a contour with downstepping level pitches is distinct from one in which the pitch slithers down from one (!)H* to the next, following the general ToBI convention of interpolating between one tonal target and the next. ToDI does not make this assumption.

5.7 Conclusion

ToDI is the most complete transcription system ever to be proposed for Dutch. It was developed in collaboration with Toni Rietveld and Jacques Terken, as well as a number of researchers in the area of prosody who took part in the Nijmegen prosody group in 1998. It is largely based on my autosegmental description of Dutch (Gussenhoven 1988, 1991, van den Berg et al. 1992), more particularly on the ‘spelled-out’ version as used in the Nijmegen synthesis-by-rule programme NIROS (Gussenhoven and Rietveld 1992).

In order to make the system accessible and learnable, an interactive course was developed for the internet in collaboration with Jacques Terken and Toni Rietveld, and designed by Ludmila Menert and Arthur Dirksen of Fluency Speech Technology in Utrecht. After a brief trial period by users of a cd-rom containing a first version, the course was made freely accessible on the Internet in July 1999 (<<http://lands.let.kun.nl/todi>>). A second edition, which includes a synthesis facility, was posted in 2004. The synthesis facility allows users to compare the intonation contour of the original speech file with that of a synthesised version, and on the basis of this comparison, adjust their initial transcription.

References

- Beckman, M.E. and Ayers, G.M. (1994), Guidelines for ToBI transcription (version 2.0, February 1994).
- Beckman, M., and Pierrehumbert, J.B. (1986), 'Intonational structure in English and Japanese', *Phonological Yearbook*, 3: 255-310.
- Berg, R. van den, Gussenhoven, C. and Rietveld, T. (1992), 'Downstep in Dutch: Implications for a model', in G. Docherty and D.R. Ladd (eds), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (Cambridge: Cambridge University Press).
- Booij, G.E. (1995), *The Phonology of Dutch*. Oxford: Oxford University Press.
- Collier, R. and 't Hart, J. (1981), *Cursus Nederlandse Intonatie*. (Leuven: Acco).
- Cruttenden, A. (1994). *Intonation*. 2nd Edition (Cambridge: Cambridge University Press).
- Féry, C. (1993), *German Intonational Patterns*. Tübingen: Niemeyer.
- Grabe, E. (1998), *Comparative intonational phonology: English and German*. PhD dissertation University of Nijmegen (Nijmegen: MPI Series in Psycholinguistics 7).
- Gussenhoven, C. (1983), 'A semantic analysis of the nuclear tones of English'. Indiana: IULC. (Included as chapter 6 in C. Gussenhoven, *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris, 1984.)
- Gussenhoven, C. (1988). Adequacy in intonation analysis: The case of Dutch. In H. van der Hulst and N. Smith (Eds.), *Advances in Nonlinear Phonology*, (pp. 95-121). Dordrecht: Foris.
- Gussenhoven, C. (1990). Tonal association domains and the prosodic hierarchy in English. In S. Ramsaran (ed.) *Studies in the Pronunciation of English*. London: Routledge. 27-37.
- Gussenhoven, C. (1991). Tone segments in the intonation of Dutch. In T.F. Shannon and J.P. Snapper (Eds.), *The Berkeley Conference on Dutch Linguistics 1989*, (pp. 139-155). Lanham MD: University Press of America.
- Gussenhoven, C. (1993). The Dutch foot and the chanted call. *Journal of Linguistics* 29, 37-63.
- Gussenhoven, C. and Rietveld, T. (1992). "A target-interpolation model for the intonation of Dutch". *ICSLP2*, pp. 1235-1238.
- Gussenhoven, C., and Rietveld, T. (2000), 'The behavior of H* and L* under variations in pitch range in Dutch rising contours'. *Language and Speech* 43: 183-203.
- Gussenhoven, C. and Bruce, G. (1999), 'Word prosody and intonation', in H. van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe* (Berlin: Mouton de Gruyter). 233-271.

- Gussenhoven, C. (2003). *Vowel duration, syllable quantity and stress in Dutch*. In K. Hanson & S. Inkelas (eds.) *The Nature of the Word. Essays in Honor of Paul Kiparsky*. Cambridge, Mass.: MIT Press. Also ROA 381.
- Grice, M. (1995), Leading tones and downstep in English. *Phonology* 12, 183-233.
- Halliday, M.A.K. (1970), *Intonation* (London: Oxford University Press).
- Hart, J. (1998), 'Intonation in Dutch', in D. Hirst and A. Di Cristo (eds), *Intonation Systems. A Survey of Twenty Languages* (Cambridge: Cambridge University Press). 96-111.
- Hart, J., Collier, René and Cohen, Anthonie (1990), *A Perceptual Study of Intonation: an Experimental Phonetic Approach* (Cambridge: Cambridge University Press).
- Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies* (Chicago: University of Chicago Press).
- Hayes, B. and Lahiri, A. (1991), 'Bengali intonational phonology'. *Natural language and Linguistic Theory* 9: 47-96.
- Kohler, K. J. (1990), 'Macro and micro F0 in the synthesis of intonation', in J. Kingston and M.E. Beckman (eds), *Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech* (Cambridge: Cambridge University Press). 115-138.
- Ladd, D.R. (1980), *The Structure of Intonational Meaning: Evidence form English* (Indiana: Indiana University Press).
- Ladd, D.R. (1983), 'Phonological features of intonational peaks'. *Language* 59: 721-759.
- Ladd, D.R. (1996), *Intonational Phonology* (Cambridge : Cambridge University Press).
- Nolan, F. and Grabe, E. (1997) 'Can ToBI transcribe intonational variation in the British Isles?', in A. Botinis, G. Kouroupetroglou and G. Caryannis (eds.), *Intonation: Theory, Models and Application. Proceedings ESCA workshop*. (Athens: ESCA and University of Athens).
- O'Connor, J.D. and Arnold, J.F. (1973), *Intonation of Colloquial English* (London: Longman).
- Pierrhumbert, J. (1980), *The phonetics and phonology of English intonation*. PhD Dissertation MIT. (Published by Garland Press, New York, 1990.)
- Post, B. (2000), *Tonal and Phrasal Structures in French Intonation* (The Hague: Thesus, Holland Academic Graphics).
- Trim, J.L.M. (1959), 'Major and minor tone groups in English'. *Le Maître Phonétique* 112: 26-29.

Visch, E. (1999), 'Stress shift', in H. van der Hulst (ed.) *Word Prosodic Systems in the Languages of Europe* (Berlin: Mouton de Gruyter). 161-232.

Zonneveld, W., Trommelen M., Jessen, M., Rice, C., Bruce, G. and Árnason, K. (1999), 'Wordstress in West Germanic and North Germanic languages', in H. van der Hulst (ed.) *Word Prosodic Systems in the Languages of Europe* (Berlin: Mouton de Gruyter). 477-603.

¹ The constituent was referred to as the Association Domain (AD) in my own work (e.g. Gussenhoven 1990), since the intonationally defined constituent does not always coincide with the Intonational Phrase as defined by other criteria. This complication is ignored here.

² The classification of the delayed peaks and the low low rise in a single morphological category having the meaning of 'significance' was rejected for English by Cruttenden (1994:117), and has found no support elsewhere. The ToDI transcription with prefixed L* thus follows Ladd's (1980, 1983) classification of 'scooped' or [+delay], which only groups 'rise-fall' and 'rise-fall-rise' of the British tradition together as versions of the 'fall' and 'fall-rise', respectively. As far as I am concerned, the jury is still out on whether the low low rise shares a semantic category with the delayed contours.

³ These tables were drawn up in consultation with Jacques Terken. I remain responsible for the interpretations given in them.