Interaction of tone and particle in the signalling of clause type in Dutch

Vincent J. van Heuven and Robert S. Kirsner
Leiden University/ULCL & UCLA

1. Introduction

1.1 Intonation

In the present paper we concentrate on one small part of the intonation system of Dutch: the options that are available to the speaker to terminate an intonation phrase (henceforth IP), a melodic pattern characteristic of a clause or even a complete short utterance.

The grammar of Dutch intonation (‘t Hart, Collier and Cohen 1990, Rietveld and van Heuven 2001:263–270) models the sentence melody of Dutch as a system of two gently declining reference lines, nominally 6 semitones (half an octave) apart, between which the pitch rises and falls in a limited number of patterns. The grammar provides three ways in which an IP may terminate: (i) on the low reference line (‘0’), (ii) on the high reference line (‘Ø’), or (iii) by executing a steep pitch rise (‘2’). The offset of rise 2 may exceed the the high reference line, specifically when the rise starts at the high reference line. The grammar then allows IPs to end at three different pitches: low, high, and extra high.

A more recent model (Gussenhoven, Rietveld and Terken 1999, Rietveld and van Heuven 2001:270–277) is constructed along the principles adopted by autosegmental intonologists, in which a sentence melody is basically a sequence of tonal targets of two types: ‘H’ (high) and ‘L’ (low). The ToDI system (Transcription of Dutch Intonation, see www.lands.kun.nl/todi or Rietveld and van Heuven 2001:399–401) is an inventory of tonal configurations for surface-level transcriptions of Dutch sentence melodies using the autosegmental H/L notation format. ToDI provides three symbols for marking IP boundaries: (i) ‘L%’, i.e., the final pitch target extends below the baseline, (ii) ‘%’, i.e., the absence of a tonal IP boundary marker, and (iii) ‘H%’, i.e., the final pitch is higher than the preceding pitch.1

Remijsen and van Heuven (1999) sought to establish the perceptual boundary
between sentence-final statement and question intonation. They did this by varying the pitch configuration on the utterance-final syllable between a fall and a steep rise in eleven perceptually equal steps. Listeners then decided for each pattern whether they perceived it as a statement or a question. At the time, we tacitly assumed that the continuum spanned just two pragmatic categories, i.e. statement versus question, and that there was no relevant intermediate category that could be interpreted as ‘non-finality’. However, Caspers (1998) suggested a functional difference between the non-tonal boundary (%) and the high boundary (H%). Her results indicate that listeners unequivocally expect the speaker to continue after the ‘H* … %’ configuration, in contradistinction to the ‘H* … H%’ pattern, for which the responses were equally divided between ‘same speaker will continue’ and ‘interlocutor will take over (with a response)’.

The present study was set up to address the question whether the continuum between the low and the high final boundary tone codes a binary contrast between finality (statement/command) and non-finality (continuation/question) or a ternary contrast between statement/command, continuation and question. It will be our working hypothesis that there is a straightforward one-to-one mapping between tonal form and pragmatic function in IP-final tone configurations, such that L% marks ‘statement’, % ‘continuation’ and H% ‘question’.

1.2 Clause typing

Dutch has lexico-syntactic means to express a range of clause types, such as statement, command, exclamation and question. Although the lexico-syntactic means are generally adequate and sufficient to express the speaker’s pragmatic intention, several — if not all — clause types are supported by appropriate intonation patterns. In fact, exceptional situations may arise where there is no lexico-syntactic differentiation between the clause types, and where the speaker’s intention can only be recovered from melodic cues. As a case in point, recent work (van Heuven and Haan 2000, Haan 2002) has demonstrated that the prosodic contrast between statement and question is, in fact, more clearly marked in the declarative question type question (which is lexico-syntactically identical to the statement) than in other question types such as the yes/no-question (which has inversion of subject and finite verb). For the purposes of the present experiment we have looked for a situation in which the three prosodic categories may serve as the only cue to a ternary choice among clause types, so that prosody will be exploited to the utmost, and the listener’s choice will not be co-determined by lexical and/or syntactic cues. Such a situation may be obtained in a V1 sentence, where the finite verb has been moved into the sentence-initial position.

In the phrase Neemt u de trein ‘Take you the train’ the lexico-syntactic information is compatible with three interpretations: 3
i. A polite imperative (Kirsner, van Heuven and Caspers 1998)
ii. A conditional subclause similar in meaning to ‘If you take the train …’
iii. A yes/no-question ‘Do you take the train?’

Which of the three readings is meant by the speaker, is only expressed through prosody. In setting up the experiment we assumed that there is no principal difference in the speech melody between a statement and a command in Dutch.4 Using a range of terminal pitch patterns on the single phrase Neem _u_ de trein, we can determine the category boundaries between command (for statement), conditional (for continuation), and question without any interfering differences in lexico-syntactic structure.

1.3 Particles

Rather than introducing uncontrolled lexico-syntactic differences, we decided to orthogonally vary the presence or absence of a particle, such that the choice of particle would systematically bias the listener against or in favor of one of the three possible readings of the stimulus.

If it is true that question and continuation are different functions carried by contrastive tonal categories, we should be able to show that the two members of the contrast respond differently to a particle that is incompatible with the pragmatic function of interrogativity. Similarly, if continuation and command are different functions carried by contrastive tonal categories, the two members of the contrast should respond differently to a particle which is incompatible with the pragmatic function of imperativity. Beginning with the latter contrast, we note that unaccented _wel_ is compatible with questions but not commands (nor with statements and conditionals). We have Neem _u_ wel _de trein! ‘Do you truly/really/actually take the train?’ but not *Neem _u_ wel _de trein_? The reason for the incompatibility lies in the fact that _wel_ is used in questions to communicate the speaker’s doubt that the state of affairs whose status is being questioned (‘taking the train’ above), is actually the case; compare Abraham’s (1984:24) example Heb je de meteorologische dienst wel _opgebeeld_? ‘DID you call up the weather bureau?’: Accordingly one might expect a pragmatic clash between the expression of doubt and the speech act of commanding, which would be intended to bring about the very state of affairs which would be doubted if _wel_ were included: cf. ?Truly take the train! versus Do you truly take the train? We predict, then, that the perceptual boundary between the question and the conditional categories will shift such that the former category will grow at the expense of the latter. Given that the particle _wel_ is also incompatible with the command category, we expect this category to shrink as well.

Conversely, the particle _maar_ ‘but, only, just’ is used to coax the hearer to carry out the action (Foolen 1995, Janssen 1995).6 As a consequence, _maar_ is compatible
only with the imperative reading of the sentence, where it mitigates the severity or harshness of the command, but does not make sense in the conditional and interrogative interpretation.\(^7\) We predict that inclusion of the particle *maar* will bias the reader in favor of the command reading, so that the command category should occupy a larger portion of the domain-final boundary continuum, at the expense of the conditional and interrogative readings.

We may conclude this introduction by summarizing the research questions that we will address:

1. Are the domain-final boundaries L\%–\%–H\% contiguous categories along a single tonal dimension?
2. Is there a one-to-one correspondence between L\% and command, \% and conditional, and H\% and question?
3. Where are the category boundaries along the continuum between (i) L\% and \% and (ii) between \% and H\%?
4. Do the category boundaries shift under the influence of particles such that the command and question categories grow when the particle is *maar* and *wel*, respectively?

2. Method

A male native speaker of standard Dutch read the following sentences (or sentence pairs) in a sound-proofed cubicle, once with and once without the particle where indicated (in parentheses):

*Neemt u (wel) de trein? Waarom bent u dan weer te laat?*
‘Do you take (part) the train? Then why are you late again?’

*Neemt u de trein, dan komt u op tijd.*
‘If you take the train, you will be in time.’

*Neemt u (maar) de trein! Anders bent u te laat.*
‘Do take (part) the train! You will be late if you do not.’

The utterances were recorded onto digital audio tape (DAT) using a Sennheiser MKH 416 unidirectional condenser microphone, transferred to computer disk (16 kHz, 16 bits), and digitally processed using the Praat speech processing software (Boersma and Weenink 1996). For the present experiment the second IP of the sentence (pair) was discarded. The particles *maar* and *wel* were excised from their original context (such that the second half of the segment preceding the particle as well as the first half of the segment following it were included) and inserted in the base utterance,\(^8\) so that three versions of the base were available: the original bare (or particle-less) version, the same base with *maar* spliced in, and the base with *wel* spliced in.
For each of the three sentences (bare, maar, wel) five intonationally different versions were generated using the PSOLA analysis-resynthesis technique (Rietveld and van Heuven 2001:379–380) implemented in the Praat software. The five versions were identical up to the H* target for the accent in the sentence-final word trein. From that point onwards the five versions diverged into two falls (moderate, steep), a level continuation, and two rises (moderate, steep).9 The terminal frequencies of the five versions were chosen to be perceptually equidistant, i.e., the difference between any two adjacent terminal frequencies was equal in terms of the ERB scale.10 The five pitch patterns are shown in Figure 1.

The 5 (versions) × 3 (particles) = 15 stimulus types were played back on-line three times in different random orders to a group of 64 native Dutch listeners who participated in the experiment as part of an introduction to Experimental Linguistics and Phonetics at Universiteit Leiden. Listeners were seated in an amphitheatre and listened to the stimuli being played to them over the public address system. Stimuli were separated by 3-s silent intervals. The series of 45 stimuli was preceded by five practice stimuli, making a set of 50 stimuli in all.11 Listeners decided for each stimulus whether they perceived it as (i) a statement, (ii) a conditional clause to be followed immediately by a main clause, or (iii) a question, with forced choice.

3. Results

Figures 2a–c (next page) present the results of the experiment, for the bare sentences, the maar sentences and the wel sentences, respectively. Each figure plots the percentage of responses ‘statement’, ‘conditional’, and ‘question’ as a function of the terminal frequency of the stimulus.
Figure 2. Percent responses ‘command,’ ‘conditional’ and ‘question’ as a function of step number of final boundary pitch. Panel A: ’bare’ sentences only; panel B: maar sentences; panel C: wel sentences. Further see text.
Figure 2a shows that for step 1 the majority of the responses are ‘command’ (some 60%), less than 40% are ‘conditional’ and the few remaining percents are for ‘question’. The distribution of the responses remains roughly the same during step 2, but at step 3 the most popular response is ‘conditional’. At step 4 ‘command’ is no longer an option for the listeners, while the percentage of ‘conditional’ responses has dropped down to 25. However, ‘question’ responses prevail (close to 80%). At step 5 we obtain almost 100% ‘question’ responses.

Figure 2a also contains two ‘cross-overs’. First, there is the cross-over from ‘command’ to ‘conditional’, which is located midway between steps 2 and 3. We have drawn a (black) single-headed arrow from the cross-over to the x-axis in order to facilitate the read-off: the cross-over is located at a stimulus value of ca. 2.6 (on the stimulus scale from 1 to 5). The second cross-over is between the ‘conditional’ and ‘question’ categories; it is located between step 3 and 4, or — as is indicated by the (gray) vertical arrow — at a stimulus value of ca. 3.5. The cross-over points identified here, i.e., of 2.6 and 3.5, are the so-called category boundaries; they are situated at a stimulus value where the listeners are at the equilibrium between the two categories on either side of the boundary.

When there is a full cross-over, i.e. from less than 25% responses to more than 75% responses for a particular category, it is possible to quantify the well-definedness, or sharpness, of the category. It is most conveniently expressed in terms of the distance along the x-axis between the (interpolated) stimulus values at which 25% and 75% responses for the category are obtained (indicated in Figure 2 by vertical dotted lines). In Figure 2a we observe just one well-defined category, viz. ‘question’. As the double-headed horizontal arrow shows, the uncertainty margin for ‘question’ is roughly 1 step-size on the x-axis.

Now let us turn to Figure 2b, which presents the results for the perception of the maar stimuli. Remember that the particle maar is compatible only with the ‘command’ interpretation of the stimulus. Accordingly, we expect the proportion of the ‘command’ response category to grow at the expense of the ‘conditional’ category, and possibly of the ‘question’ category as well. At stimulus step 1 the number of ‘command’ judgments is 75%, which is some 15 percentage points more than in the bare stimuli. Moreover, the category boundary between ‘command’ and ‘conditional’ has been pushed up the continuum to ca. 3.2 — which value is .6 of a step up the continuum. The second category boundary, between ‘conditional’ and ‘question’, has moved by a similar amount, i.e. from 3.5 to 4.2. As a result, the sharpness of the ‘question’ category has not changed: its uncertainty margin is still 1 step.

Figure 2c presents the results for the wel sentences, for which we expect the reverse effect to take place. Since wel is compatible with ‘question’ but clashes with the ‘command’ and ‘conditional’, the ‘question’ category should grow, or, phrased differently, the category boundaries should shift to lower values along the x-axis. Let us look at the results. Note that the boundary between ‘command’ and ‘conditional’
has shifted to a lower value on the x-axis of 2.2, i.e., a drop of .4 relative to the particle-less edition of the stimuli, which is as predicted. However, the second boundary, that between 'conditional' and 'question' has remained stable. Also, the steepness of the question cross-over has not been affected by the particle: its uncertainty margin is still in the order of 1 step-size along the continuum.

4. Conclusions and discussion

Our first two research questions asked if the tone range between the low (L%) and the high (H%) edge tone has an intermediate formal category (%) such that the three categories constitute a contiguous functional/formal division of one tonal continuum. The results of the experiment indicate that this is indeed the case. Figures 2a–c show that the listeners neatly divide up the continuum in three response categories, such that a specific, separate communicative function is associated with each formal category. L% is associated with 'command', % with 'conditional', and H% with 'question'. Each response category peaks at a specific stimulus value in a non-random fashion (see below). Although this result is not surprising, it is important to note that existing accounts of Dutch intonation have neglected to make the ternary formal and functional division of the edge tone continuum explicit. Earlier experimental evidence that the non-low part of the edge tone continuum should be split up between two separate formal categories, each with an identifiable communicative function, was circumstantial (Caspers 1998, van Heuven and Haan 2000, Haan 2002).

Let us next consider the phonetic characterization of the three formal edge tone categories. We may split this issue up into two sub-questions: (i) what are the prototypical exemplars of each category, and (ii) where are the boundaries between contiguous categories? The answer to (i) is straightforward. It is always the case that the contour with the lowest terminal pitch (step 1) is most strongly felt as a command. The contour with the highest terminal pitch (step 5) invariably comes out as the prototypical question. Continuation (or non-finiteness, here 'conditional') is typically associated with step 3 on the continuum but the association varies under the influence of the particle (see below).

The boundaries between the three tonal categories (L%, %, H%) can best be observed in the bare stimuli, i.e., in the absence of possibly interfering particles. The results reveal that both the steeper and the shallower fall lead to a preponderance of 'command' responses. To us this would indicate that a relatively gentle domain-final fall is all that it takes for the speaker to signal the low boundary tone. Similarly, both the gentle and the steeper rise after H* are sufficient means to communicate interrogativity, since both these stimulus types lead to a strong majority of 'question' responses. The recipe for %, i.e., the absence of a tonally marked IP boundary, should then be: after the last accent in the IP, the pitch should remain essentially flat.
Crucially, however, the boundaries between the three tonal categories are not fixed but interact, i.e., move up and down the pitch scale (while preserving the same linear sequence), under the influence of particles that predictably affect the interpretation of the utterance. The categories do not expand and shrink in an arbitrary, random fashion but expand to ‘invade’ some of the space occupied by the ‘impossible’ category (or categories) that are pragmatically incompatible with a given particle. When the ‘question’ category shrinks, due to the interaction with maar — which blocks the ‘question’ interpretation — both category boundaries, the L% ~ % boundary and the % ~ H% boundary, shift to higher values. That is, both the non-question category of ‘command’ and the non-question category of ‘conditional’ invade the space occupied in bare, particle-less sentences by the question category. Under the influence of wel, something like a mirror-image situation arises, but with a curious asymmetry: the semantically incompatible command category shrinks, i.e., shifts downwards, in the opposite direction, to a lower value, but question category does not expand, does not shift downwards to match the shift of the command category. The question category stays put, leaving us with a de-facto enlarged conditional category. Presumably this asymmetry can be explained by the necessity of having at least some degree of final rise (steps 4 and 5) for a question interpretation to be at all possible, so that while the question category could in principle expand into space held by the conditional category, it does not. This asymmetry, too, would seem to be consistent with our hypothesis that the relevant pragmatic division of the intonational space is not binary but ternary. The three pragmatic categories never overlap, are never superimposed, but preserve their distinctness, even though the degree of distinctness differs when one extreme or the other of the step-continuum (and never the middle) is occupied by a semantically incompatible, ‘impossible’ combination of pragmatic particle and sentence-type.

Finally, the results show that full cross-overs between categories are obtained only between continuation and question. In all three figures this cross-over is relatively steep, with an uncertainty margin of one step-size along the continuum. Moreover, the steepness was not affected by boundary shifts as a result of interaction with a particle. The boundary between command (finality) and conditional (non-finality), however, is relatively poorly defined: the cross-overs are incomplete, and in the one case where it is almost complete, it is shallow. The relatively poor separation between the response categories ‘command’ and ‘conditional’ might, at face value, be interpreted as counterevidence to our claim that the edge tone continuum has three rather than two formal categories (with corresponding functions). In order to settle the issue definitively, we would really need a new experiment, in which the same stimulus materials that were used in the present study, are offered to (preferably the same) listeners three more times. In one condition the listeners will be instructed to identify the stimuli in terms of a binary
division: ‘command’ versus ‘non-command’, in the second condition the response categories will be ‘question’ versus ‘non-question’, and in the last condition the listeners would choose between ‘command’ and ‘question’. The location of the category boundaries and the steepness of their cross-overs could then be compared with the results of the present experiment. If there were truly no intermediate ‘continuation’ category, then the cross-over from ‘command’ to ‘question’ should be poorly defined, while the cross-over between ‘command’ versus ‘non-command’ and that between question ‘versus ‘non-question’ should be steep.

Notes

1. The ‘%’ sign following the tone letter (as in L%, H%) denotes a domain-final boundary; domain-initial boundaries are coded by a tone letter preceded by the ‘%’ sign (as in %L, %H). A ‘%’ sign unaccompanied by a tone letter may only occur in domain-final positions, where it is phonetically coded by pre-boundary lengthening only.

2. In fact, Kirsner and van Heuven (1996) suggested a single abstract meaning for the non-low tonal category: ‘appeal (by the speaker to the hearer)’, asking for the hearer’s continued attention or for a verbal response to a question or a non-verbal compliance with a request.

3. An anonymous referee has suggested that a sequence like Neemt u de trein might in addition be interpretable as a topic-drop-sentence (e.g. [Dan/Daar] wu b u de trein ‘[Then/There] you take the train’, analogous to Doen we! ‘We’ll do [it]’ or Weet ik? ‘[That] I know’. Although this added interpretation (with a ‘deleted’ element) is theoretically possible, we believe that it was highly unlikely under the controlled conditions of the experiment. Moreover, none of the experimental subjects ever volunteered a remark to the effect that they had missed a fourth response category.

4. This position does not exclude the possibility that statement and imperative are subtly different in their paralinguistic use of prosody. For instance, the overall pitch of the imperative may be lower, and it may be said with greater loudness and larger/higher pitch excursions on the accented syllables. This does not invalidate our claim that both statements and interrogatives are coded by the L% terminal boundary.

5. If it were simply a question of persuading the hearer to take the train as opposed to not taking it, one would theoretically have to use accented, non-modal wel.

6. Cf. van der Wouden’s (1998) gloss of the effect of maar as ‘feel free to’ and the Van Dale (Martin and Tops 1986) gloss of ‘go ahead and’; as in Neemt u maar de trein ‘Go ahead and take the train’.

7. The anonymous referee has suggested that, in certain cases, maar occurs in certain V1-utterances which can be said to have the structure of questions in that they are pronounced with ‘question intonation’. Nevertheless, the fact that a sentence such as Zullen we maar naar bed gaan? ‘Shall we be going along to bed?’ (cited in Van der Wouden 2002:38) functions pragmatically not as a genuine question but as a weak command or suggestion would seem to only prove our point about the incompatibility of maar with questions. In the context of our experiment, we would argue that the interpretation of Neemt u maar de trein? as a genuine question, hence as communicating something as unusual as ‘Do you feel free to take the train?’ or ‘Are you going ahead and taking the train?’, was extremely unlikely.
8. In fact, the bare version of the conditional was selected as the base form from which all further stimuli would be generated. It was felt that the intermediate final rise of the conditional would be easiest to convert to either an L% or H%.

9. The falls on trein are implementations of ‘H* L%’ tone configurations. The steep part of the fall is due to the H element, the shallower trailing part of the fall is the implementation of the edge tone L%. In the flat and rising patterns there is no tone between H* and the (H)% edge tone.

10. The ERB scale (Equivalent Rectangular Bandwidth) is currently held to be the most satisfactory psychophysical conversion for pitch intervals in human speech (Rietveld and van Heuven 2001).

11. The subjects also participated in a second experiment, in which they were asked to rate the relative degree of coherence of the nine combinations of the three first sentences/clauses listed at the beginning of this section with the three second sentences/clauses (each combination with the five pitch patterns on the first IP, yielding a total of 45 stimulus types), e.g. Neemt u maar de trein, dan komt u op tijd. Space limitations preclude discussion of this experiment here.

12. This state of affairs involving intonation is perhaps somewhat reminiscent of Martinet’s (1955) notion of a ‘margin of security’ between segmental phonemes.

References


Language and Speech 41, 375–398.


