

Nederlandse Taalkunde  
www.nederlandsetaalkunde.nl  
Uitgave: Amsterdam University Press

# Prosody and sentence type in Dutch<sup>1</sup>

Vincent J. van Heuven

NT 22 (1): 3–29

DOI: 10.5117/NEDTAA2017.1.HEUV

BELANGRIJKE OPMERKING. Dit artikel bevat een digitale bijlage met de volgende DOI:  
10.6084/M9.FIGSHARE.4689667.V2

## Abstract

This article summarizes earlier research done on the prosodic marking of interrogativity and imperatives in Dutch on the basis of recorded speech from male and female speakers. The first part of this article compares statements (ST) and three types of question. The form of questions may differ in various respects from statements: Wh-questions (WH) have a question word in initial position and exhibit subject-verb inversion, yes/no-questions (YN) have inversion only, while declarative questions (DE) have the same structure as ST. Our functional hypothesis that the intensity of interrogativity marking through prosody counterbalances the degree of syntactic marking in the order  $ST < WH < YN < DE$  is confirmed both by an analysis of the phonological choices made and by the details of the phonetic implementation of the melodies. The second part of the article compares statements and imperatives. All sentences were produced with three different attitudes: neutral, friendly and authoritarian. Our hypothesis that imperatives are prosodically marked in the same way as authoritarian statements seems to be confirmed, as an analysis of tone transcriptions reveals no systematic differences between them at the linguistic level. The difference is only weakly marked in the prosody at the phonetic level, in a way that suggests that the speaker raises his/her voice in imperatives (higher overall pitch, longer duration, greater intensity), which is also the way in which the authoritarian style of speaking differs from the other attitudes.

**Keywords:** Dutch, prosody, sentence type, interrogative, imperative, speaker attitude, tone transcription, phonetic correlates, pitch, intensity, duration

<sup>1</sup> I thank the reviewers and especially the editors of the present volume for their meticulous reading of earlier versions of this article and many valuable suggestions for improvement.

# 1 Introduction

## 1.1 What is prosody?

In the study of speech and language it is customary to analyse a sentence as a sequence of discrete time-ordered units often referred to as segments. A spoken sentence is analysed as a sequence of morphemes, i.e., minimal sequences of sound with an identifiable meaning associated with them, while each morpheme is decomposed into a sequence of vowels and consonants – which are seen as the shortest building blocks from which spoken sentences are constructed. At the same time it is readily acknowledged that words and sentences cannot be described as just a sequence of segments. In fact, the same sequence of vowels and consonants may constitute different words with different meanings (and the same sequence of words may make up different sentences) depending on the prosody. Prosody is defined, then, as the ensemble of properties characterising words and sentences that cannot be predicted from the mere sequence of smaller sound segments.<sup>2</sup> For instance, Dutch has two different words with unrelated meanings that are composed of the same five sounds in the same order: /χɑ:vɔn/. If the first syllable is pronounced stronger than the second, the meaning is ‘list of saints’ or ‘round song’ (spelled *canon*) but with the emphasis on the second syllable the word means ‘large gun’ (spelled *kanon*; for more discussion of the definition of prosody see Van Heuven 1994, 2008; Van Heuven & Sluijter 1996; Nootboom 1997; Rietveld & Van Heuven 2016, chapter 12).

It is customary to differentiate between word prosody and sentence prosody. There are basically two types of word-prosodic systems in the languages of the world, i.e., stress languages and tone languages. The communicative function of tone would be to help differentiate between words in the lexicon. For instance, in Mandarin Chinese, the same syllable /ma/ may signal five different and unrelated meanings, depending on the tone (e.g. Van Heuven 2008). Stress may also help to differentiate between words (as in *canon* ~ *kanon* above) but only in languages in which the location of the stressed syllable within the word is unpredictable. In languages with fixed stress, such as in Hungarian where stress is always word-initial, stress may signal word boundaries. The communicative functions of

2 The word *prosody* derives from Greek *pros* ‘with’ and *oidein* ‘sing’, i.e., it refers to the musical accompaniment of sung words. Somewhat confusingly, the Latin loan translation *accent* (= *ad* + *cantus* ‘with the song’) does not refer to all aspects of prosody but is restricted to the phonetic realisation of emphasis only.

sentence prosody are many. By acceleration (initial strengthening) and deceleration (pre-boundary lengthening), and pitch changes the speaker may signal which sequence of words should be analysed as a prosodic domain – with different markings depending on the hierarchical depth of the domain. Within such prosodic domains one word is typically presented as communicatively more important than the others through sentence stress – typically marked by a conspicuous change in vocal pitch on the syllable within the word that also has word stress (see e.g. Van Heuven 2014a). Different melodies may then further qualify the information status of referents relative to the background shared by speaker and hearer. Prosody also makes an important contribution to the expression of affect (an umbrella term covering emotions and attitudes on the part of the speaker). The last function of prosody is to qualify a sequence of words, be it a sentence or a clause within a sentence, as a specific type. This latter function is often referred to as sentence typing or clause typing. Well-known sentence types are statement, question, exclamation, command, and negation. On a lower level, there are at least two clause types, i.e., clauses that are sentence final (finality) and non-final clauses, which typically contain a prosodic continuation marker.

There has been considerable debate on the relative importance of segmental structure versus that of prosody. In written language, prosody is hardly reflected while most orthographies provide a fairly precise and exhaustive coding of the identity of the vowels and consonants that make up a sentence. Word and sentence stress is normally not marked in written language at all, while only the deepest prosodic boundaries (corresponding approximately with syntactic paragraph, sentence and clause boundaries) are represented in spelling. Prosody is almost completely redundant given the segmental make up of the words in sequence. If all prosodic information is eliminated from a spoken Dutch (or English) sentence, for instance using speech synthesis based exclusively on sounds excerpted from fully stressed syllables and omitting any pitch inflection or adjustment of segment durations, the result is perfectly intelligible. However, if we eliminate all segmental information from an utterance while fully maintaining prosodic information, no intelligibility remains.<sup>3</sup> Even in tone languages such as Mandarin, intelligibility remains at ceiling level when prosody (including word tones) is eliminated (Qin 2007, Van Heuven 2008, Wang et al. 2011). A different state of affairs emerges if we do not just eliminate an

3 These effects are demonstrated here: <https://figshare.com/s/7abob2461206b42f53aa> (DOI: 10.6084/m9.figshare.4689667.v2).

information source from the speech signal but create an artificial conflict between segmental and prosodic information. As long as the speech quality is good, so that vowels and consonants are correctly identified by the hearer, a mismatch with the prosody is readily spotted by the native listener but subsequently ignored in the word recognition and sentence comprehension processes. However, if the segmental quality is degraded, e.g. by ambient noise, poor transmission lines or foreign accent, segmental information is re-interpreted so as to match the more robust, noise-resistant prosodic information. This is generally a helpful strategy but it backfires when the prosody is wrong.

As a case in point, Bansal (1966) showed that native listeners of British English reinterpret segments so as to fit the incorrectly perceived rhythmic pattern in English pronounced by Indian speakers. To the British-English ear the stress in the words *character* and *written* is perceived on the second syllable when pronounced by an Indian speaker of English. Given that the segmental pronunciation of Indian English strongly deviates from the British norm, it makes sense for the British listener to reinterpret the information as *director* and *retain*, respectively (errors reported by Bansal).

Similarly, at the sentence level, removing prosodic information from utterances is largely inconsequential. However, when a conflict is created between the syntactic-semantic structure of the sentence and the prosody, listeners tend to attach more weight to the prosodic information if the speech quality is poor (e.g. Wingfield 1975).

The purpose of this article is to give a summary of work that I did with a number of co-workers on the prosodic marking of sentence types in Dutch. I will concentrate on two lines of research. After a short introduction in section 2, the first issue we address is how the difference between statements and questions is signalled by prosody (section 3). The second issue is whether Dutch has a way of prosodically marking the difference between statements and commands (section 4).

## 2 Prosody and sentence type in Dutch

Many languages employ dedicated syntactic structures to differentiate between sentence types. In Dutch, questions (interrogatives) may be signalled by inversion of subject and finite verb (*Kom jij vanavond?* ‘Come you tonight?’), or by having the sentence begin with a question word (*Wie komt er vanavond?* ‘Who comes there tonight?’. *Wanneer kom jij?* ‘When come you?’). Commands (imperatives) use the shortest possible verb form

and omit the subject/addressee (or use it in the vocative: *Val dood, man!* 'Drop dead, man!'). Other sentence types exist as well, such as exclamatives and negations, but these will not be part of the present research.

It seems an attractive idea that each of these sentence types should be redundantly signalled by prosodic means. Questions are almost universally marked by some high-pitched element in the sentence melody which would typically be absent from statements. Similarly, one finds claims in the literature that commands are characterised by a special melody or tone of voice. Some languages are even claimed to possess a special melody that expresses negation. Orthographies have special (sentence-final) markers to differentiate these prosodic sentence types: the question mark '?' for interrogatives and the exclamation mark '!' at the end of imperatives and exclamatives. The full stop '.' would signal the end of a statement. In everyday speech communication, however, the various melodies can be combined with each of the syntactic sentence types: there is no one-to-one relationship between syntax and prosody. A more adequate analysis, then, would be to assume that there is no compulsory coupling between syntactic and prosodic sentence types. Rather, the various syntactic types have a meaning of their own, while the prosodic types may have different meanings. In this way, one may meet with questions that have the prosody of statements, or vice versa. For instance, a wh-question with the typical melody of a question expresses a rather different meaning than the same structure produced with statement melody. When a wh-question such as *Waarom eet je geen vlees?* 'Why eat you no meat?' ends on a high pitch, typical of questions, the speaker would be sincerely interested in receiving an answer, whereas the same sequence of words produced with falling intonation would rather be interpreted as a reproach, requiring no answer from the hearer. Similarly, an (affirmative) statement spoken with question intonation, a so-called declarative question, is not a prototypical question; using this type of hybrid structure the speaker indicates to the hearer that he would like to receive a response but at the same suggests that the answer would most likely be 'yes'.<sup>4</sup>

4 Haan and Van Heuven (2003) provide a detailed experimental study that shows that the regular yes/no-question elicits 'yes' versus 'no' responses in roughly equal proportions, whereas the declarative question version of the same structure was strongly biased for 'yes' responses.

### 3 Prosody of Dutch questions

We examined the phonetic markings of three basic question types in Dutch and compared these with the properties of (corresponding) statements as summarized in (1a-d).

- |                                    |  |
|------------------------------------|--|
| (1) a. ST: Statement               | Jij bent morgen ziek.<br>'You are tomorrow ill.' |
| b. YN: Polarity or yes/no-question | Ben jij morgen ziek?<br>'Are you tomorrow ill?'  |
| c. WH: Information or wh-question  | Wanneer ben jij ziek?<br>'When are you ill?'     |
| d. DE: Declarative question        | Jij bent morgen ziek?<br>'You are tomorrow ill?' |

The work was done in a laboratory setting, in which phonetically naïve speakers of Dutch were instructed to read out pairs of sentences as if they were actors in a radio play.<sup>5</sup> In the pairs one sentence provided the context while the other sentence was the target. As a control measure, the target sentences were also spoken in isolation.

Using the statement version (1a) as the unmarked structure, the yes/no-question type (1b) differs from the former in just one syntactic marker, i.e., the inversion of subject and finite. The wh-question type (1c) we used differs in two lexico-syntactic characteristics from the corresponding statement, viz. the question word, and the inversion of subject and finite. The declarative question type (DE), however, is identical to the statement in lexis and surface syntax. We therefore expect the need for prosodic marking of interrogativity to be strongest for the declarative question, intermediate for the yes/no-question and least for the wh-question type, i.e., DE > YN > WH > ST. This basic, functionally inspired, hypothesis was tested in the experiment.

Interrogativity is prosodically marked in the large majority of the world's languages by the presence of some high-pitched element that is absent from the statement (Lindsey 1985). The high pitch may be caused by the choice of phonological tone pattern such that H elements replace L elements, especially towards or at the end of the sentence. Alternatively, the high pitch may be caused by phonetic implementation of tones where-

5 Much of what was found in the laboratory study was later corroborated by an analysis of recordings of spontaneous speech (see Van Heuven et al. 1998).

by these are realised on a higher pitch. These two possibilities, which are not mutually exclusive, will be tested in the experiment.

### 3.1 Method

Haan et al. (1997) recorded statements and questions in a controlled study from five male and five female Dutch speakers. Sentences were statements, yes/no-questions, wh-questions and declarative questions of two lexically different base structures, as exemplified in Table 1.

**Table 1** Target and context sentences used in Haan et al. (1997). Context sentences were present or absent, and if present preceded or followed the target sentence. Both context sentences were combined with all four target sentences.

Type	Target sentence	Type	Context sentence
ST	<i>Renée heeft nog vlees over.</i> 'Renée has some meat left.'	ST	<i>Onze poes moet wat eten hebben.</i> 'Our cat must some food have.'
WH	<i>Wat heeft Renée nog voor vlees over?</i> 'What has Renée still of meat left?'		
YN	<i>Heeft Renée nog wat vlees over</i> 'Has Renée still some meat left?'	YN	<i>Wil de poes nog wat eten hebben?</i> 'Wants the cat still some food have?'
DE	<i>Renée heeft nog vlees over?</i> 'Renée has some meat left?'		
ST	<i>Marina wil haar Mandoline verkopen</i> 'Marina wants her mandolin sell.'	ST	<i>Er is donderdag weer een rommelmarkt.</i> 'There is Thursday again a jumble sale.'
WH	<i>Wanneer wil Marina haar mandoline verkopen?</i> 'When wants Marina her mandolin sell?'		
YN	<i>Wil Marina haar Mandoline verkopen?</i> 'Wants Marina her mandolin sell?'	YN	<i>Is er donderdag weer een rommelmarkt?</i> 'Is there Thursday again a jumble sale?'
DE	<i>Marina wil haar mandoline verkopen?</i> 'Marina wants her mandolin sell.'		

Omitting further experimental detail (see Haan 2001 for a full account), we computed the fundamental frequency, as the acoustic correlate of vocal pitch, by the method of subharmonic summation (Hermes 1988) at 10-ms intervals (using Praat speech processing software, Boersma & Weenink 1996) and expressed the pitches in ERB-units.<sup>6</sup> We then segmented off the

6 The Equivalent Rectangular Bandwidth was held at the time to be the psychophysically most realistic way of expressing pitch intervals in sentence melodies. More recently ERB has been shown to be the most adequate scale to capture the perceived prominence of pitch movements generating sentence stress. When we are merely interested in the melodic size of a pitch change, the musical interval scale (in semitones) is preferred (Nolan 2003). For formulae to convert frequencies in hertz (Hz) to either ERB units or semitones, see Rietveld & Van Heuven 2016: 407-408.

final boundary tone in each target utterance, which could be either L% (final lowering) or H% (final rise). The pitch interval between the beginning and end of the boundary tone was measured.

### 3.2 Analysis and results

Although descriptions of Dutch sentence melody generally assume an overall downtrend of the pitch over the course of the sentence (e.g. 't Hart et al. 1990, Van den Berg et al. 1991), there were indications in the literature that questions might be characterised by a flatter downtrend, or even an uptrend in the pitch contour (e.g. Thorsen 1980, Gooskens & Van Heuven 1995). For this reason, in the next stage of the analysis we fitted a regression line through the pitch contour over the body of the utterance, i.e., from the onset to the beginning of the final boundary tone, as shown in Figure 1 (i.e. excluding the final boundary tone, so that a final rise or final lowering had no influence on the slope of the pitch trend).

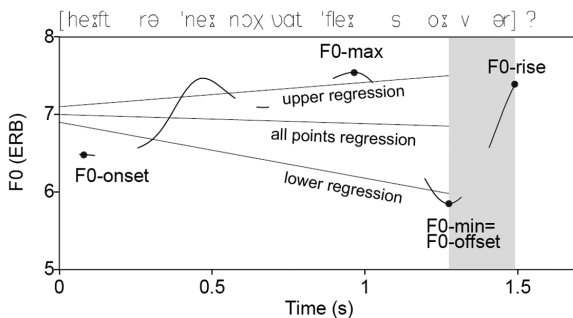


Figure 1 Fundamental frequency (ERB) of the utterance 'Heeft Renée nog wat vlees over?' The shaded segment marks the final boundary rise. Three regression lines are shown (see text for explanation, from Haan et al. 1997).

This regression line divides the pitches in an upper and a lower half while observing the overall pitch trend in the body of the utterance. However, it soon transpired that a single overall trend line does not do justice to what can be seen in the pitch curves of statements and the various question types in Dutch. Typically, the size of the pitch changes marking the sentence stresses on the question word (if present) and on the subject and object nouns differs substantially depending on the question type. It seemed better, therefore, to model the overall shape of the melodies not by two parallel (lower and upper) trend lines but to allow for converging or diverging trend lines. This was achieved by fitting new regression lines to the lower and upper half of the pitch points separately. The mean distance



(in ERB) between the upper and lower trend lines, and the slope of each trend line (in ERB per second) serve as parameters that characterize differences in the pitch spans between the sentence types.

Figure 2 shows that ST has the largest mean pitch span (1.17 ERB), while the upper and lower trend lines run parallel with a declination of roughly .5 ERB. WH is characterised by a narrow pitch span, with parallel trend lines starting at a high pitch and declining rather steeply at a rate of just over 1 ERB/s per second. With the possible exception of the high onset pitch (which, however, may well be caused by a sentence stress on the question word), the WH type bears a strong resemblance to ST. YN differs from ST in that the upper and lower trend lines no longer run parallel but diverge from beginning to end of the body of the utterance. The slope of the lower trend line is quite similar to its counterpart in ST but its overall course is elevated by about .5 ERB. Moreover, the upper trend line now shows moderate inclination (rising pitch) rather than declination. These characteristics make the YN type differ clearly from the ST melody. The pitch span of DE is basically the same as that of YN but both the lower and upper trend lines show rising pitch – as if the entire contour were rotated counter-clockwise around a central pivot point. As a result, DE differs even more strongly from ST than YN while the basic similarity in the overall configuration seems to bear out that these two question types share one functional characteristic in that both require the hearer to respond with either *yes* or *no* (but see also note 4).

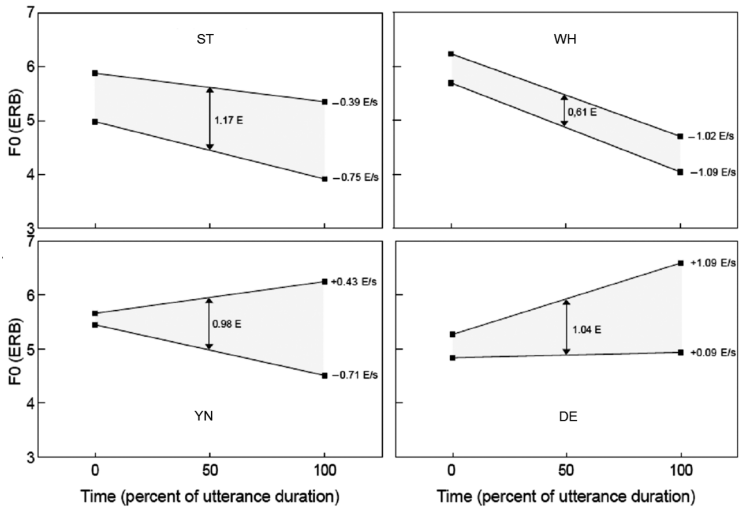


Figure 2 Stylised and time-normalised pitch span of statement and three types of question melody in Dutch (adapted from Van Heuven & Haan 2000)

In order to understand the differences in width of the pitch span and in orientation of the trend lines, we performed a more detailed stylization of the 800 utterances in the materials. Utterances contained two words that could receive sentence stress, i.e., the subject and the object nouns. In the case of WH the sentence-initial question word was a third potential landing site for sentence stress. Each potential sentence stress was modelled with three pitch points (henceforth ‘pivots’), as illustrated in Figure 3 (for details see Van Heuven & Haan 2000, Haan 2001). Figure 3 presents the mean locations of the six (or nine, in the case of WH) pivots in the frequency-time domain.<sup>7</sup>

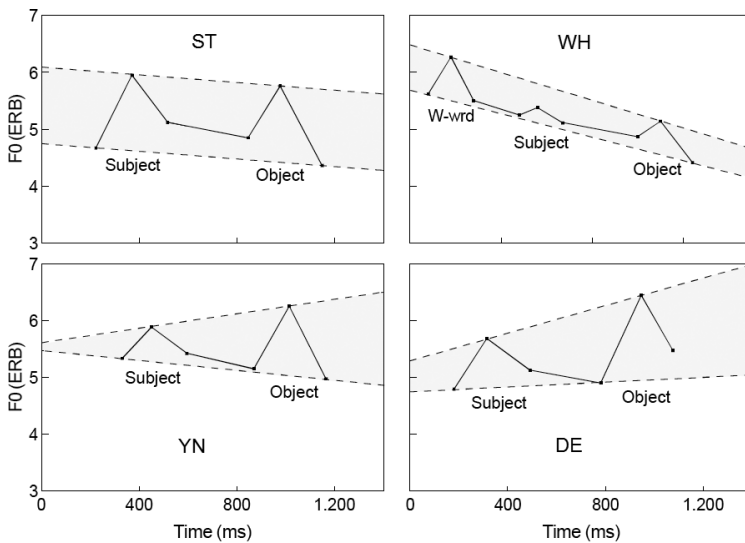


Figure 3 Stylised pitch contours of four sentence types (200 tokens per panel) in frequency (ERB) by time (ms) plots. Upper trend lines were drawn by connecting the two highest pivots in the utterance; the lower trend line connects the first low with the last low pivot (second but last for declarative questions).

Figure 3 shows that the lower trend lines in Figure 2 quite closely match the lower pivot points found in the stylised pitch contours of the four sentence types. Moreover, the upper trend lines in Figure 2 are fully explained by the difference in size of the pitch movements marking the first

7 In Figure 3 durations are not normalised but expressed in actual time units. In later research Van Heuven and Van Zanten (2005) showed that question versions of sentences tend to be spoken at a faster rate than the corresponding statements, not only in our Dutch data but also, and even more convincingly, in materials for (Orkney) English and (Manado) Malay.

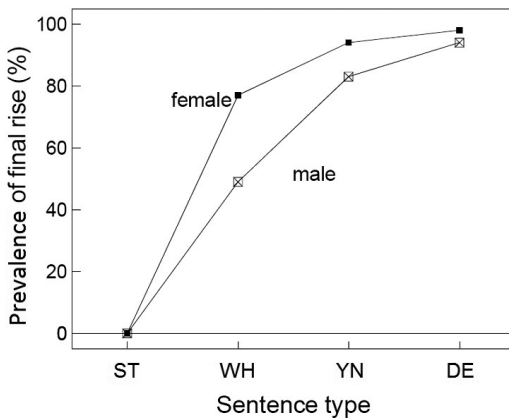
and last sentence stresses in each utterance. Table 2 summarizes the differences between the four sentence types analysed here.

**Table 2** Characterisation of four sentence types in terms question-marking phonetic prevalence

Phonetic question markers	Sentence type			
	ST	WH	YN	DE
Lower trend line rising?				+
Upper trend line rising?			+	+
Last sentence stress has highest pitch?			+	+
Total	0	0	2	3

Table 2 shows that it is generally difficult to phonetically discriminate between ST and WH melodies in the body of the sentence, i.e. excluding the marker that might be in the final boundary tone. This would tie in with the idea that WH is sufficiently distinct from ST by lexico-syntactic means: there is no need for additional prosodic marking. It can also be seen in Table 2 that the number of question-marking prosodic properties is larger for DE than for YN. This finding, too, matches the predictions of our functional hypothesis: since DE does not differ from ST, the prosody has to do more work.

Finally, let us consider the effect of sentence type on the incidence of the use of a sentence-final pitch rise (or: high boundary tone H%) in the materials collected by Haan et al. (1997). Figure 4 presents the percentage of utterances that were found to end in H%, broken down by the four sentence types separately for male and female speakers.






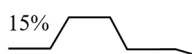

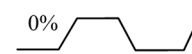

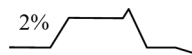




*Figure 4* Percentage of high boundary tones found at the end of four sentence types in Dutch, broken down by gender of speaker

Again, the prediction of the functional hypothesis appears to be confirmed. There is never an H% at the end of statements, while it is nearly always found after declarative questions (96%). The H% is found less often after yes/no questions (89%) and even less often at the end of a wh-question (65%).<sup>8</sup>

Table 3 presents a tone transcription of the 800 utterances in our materials, using the ToDI (Tones of Dutch Intonation, Gussenhoven et al. 2015, Rietveld & van Heuven 2016: 430-433) transcription system augmented with an upstep symbol '^', which has the opposite effect of the downstep (') and raises the pitch target of a non-initial sentence stress.

**Table 3** Melodies and transcription of 200 statements and the corresponding 200 declarative questions. Percentages per sentence type do not add up to 100. The complement is distributed over infrequent patterns.

Statements		Declarative Questions	
 68%  4%  9%  6%	 7%  15%  23%  0%	 0%  2%  38%  15%	

Limiting the discussion to just the comparison of statements and their declarative question counterparts (i.e., the question type with strongest prosodic interrogativity markers), we find the following distribution of tones: (i) statements always end in L% (and never in H%), whereas declarative questions exclusively end in H%, (ii) the last sentence stress in the statement is realised as either H\* (72%) or a downstepped !H\* (22%), while this is the upstepped ^H\* in 76% of the declarative questions (for a complete survey of tones in statements and declarative questions see Table 3 or Haan 2001). Clearly, then, the declarative question has one or more H tone elements that are absent from the statement.

8 The research also investigated the role of gender in question marking. An interaction between sentence type and gender was predicted and found, in that women tend to use the H% boundary tone more often than men. See Haan & Van Heuven (1999) for statistical analysis and discussion of this phenomenon.

### 3.3 Conclusion and discussion

The experimental evidence presented above supports the functional hypothesis that the phonetic marking of question types should be more clearly pronounced as the number of lexical and/or syntactic cues to interrogativity is smaller. Confirmation is seen not only in terms of the prevalence of the H% sentence-final prosodic question marker but also by more subtle warning signals issued earlier on in the time course of the utterance, i.e., the slope of the lower declination line (lower trend line) and in the difference in the size of the pitch movement (or height of the pitch target) between the first and the last sentence stress in the utterance.<sup>9</sup>

I end this section with a few general remarks. The first is that, according to the results presented above, Dutch behaves very much like the vast majority of the languages in the world in that interrogativity is marked by some high-pitched element that is not found in statements. The near-universality of high pitch in questions has been explained as a remnant from earlier stages in the evolution of mankind (Ohala 1981, Gussenhoven 2004, Van Heuven 2014b). By making high-pitched sounds one animal may signal submissiveness to a dominant member of the same species, in order to avoid bloodshed during a confrontation. High-pitched noises are associated with small animals, typically with the sounds produced by the young of the species, which should not be harmed by the adults. By extension, this so-called frequency code is still used by humans when asking a question. Asking for information puts the speaker in a position of weakness, in which adopting submissive behaviour (i.e., producing high pitched sounds) would be an appropriate strategy.

The second, and concluding, remark is that the H% tone should not be seen as an exclusive question marker in Dutch. Dutch distinguishes between a low and a high boundary tone in a categorical fashion. The low boundary tone L% signals the meaning of finality: end of sentence. The complementary boundary tone H% signals the opposite: the preceding unit is not finished. Instead, the speaker asks the hearer to attend to the continuation of the message, or – in the case of a question – to advance information asked for by the speaker. Crucially, the contrast between L% and H% is perceived categorically, i.e., with a sharply defined cross-over between the categories in an identification task combined with a local

9 In perception experiments we showed that Dutch native listeners use the early-warning interrogativity signals, and in fact discriminate statements from (declarative) questions quite effectively as soon as the last sentence stress is heard, i.e., one word before the end of the sentence and before hearing the final pitch rise marking interrogativity (Van Heuven & Haan 2000, 2002).

peak in the discrimination function at the category boundary that cannot be statistically predicted from the identification scores. The contrast within the H% rising boundary tones signalling non-finality versus question is continuous rather than categorical: the cross-over is gradual, and is not matched by a local discrimination peak (Van Heuven & Kirsner 2004). The upshot of this is that the contrast between L% (finality) and H% (non-finality) is linguistic in nature but that the further division within the H% category between 'continuation' and 'question' is paralinguistic (or phonetic) and should be interpreted as a continuous difference in degree of urgency ('appeal') on the part of the speaker: presumably, asking a response from the hearer presupposes a greater favour than asking for the listener's continued attention.

## 4 Prosody of Dutch commands

### 4.1 Introduction

The melody of commands in Dutch has received attention in the literature as far back as Quittart (1925). He claims that a falling intonation pattern is the rule for commands and imperatives. Since falling intonation is also the default melody of statements, this suggests that imperatives and statements basically share the same melody. Quittart also observes that a rising melody towards the end of an imperative (no pitch fall after the last sentence stress) makes the command more forceful, while a rise-fall-rise pattern (presumably a H\*L sentence stress on the last content word) followed by an H% boundary tone mitigates the command to a friendly or even condescending warning. Kirsner et al. (1998) presented experimental results which would appear to confirm Quittart's observations. It was shown that imperatives with only an H\*L sentence stress were perceived by Dutch listeners as more forceful commands than imperatives with either just a rise (%L H\*) or just a fall (%H L\*). They also showed that imperatives ending in the high boundary tone H% were perceived as weaker commands than the same patterns ending in the low boundary tone L%. Van Heuven & Kirsner (1999) showed that both raising the pitch register and reducing the size of the pitch movements on sentence stresses render imperatives less forceful.

Though the above findings tell us what not to do when we want to issue a forceful command, they do not tell us what the typical prosodic shape of imperatives (and commands in general) would be. In this article I will present the results of an experiment that was set up to shed light on

precisely this issue. In the study we recorded speakers who were asked to produce imperatives and statements with either a neutral, authoritarian or friendly tone of voice.

We expect commands to be issued with a loud voice – as a sign of physical strength, leadership and authority. Loud voices, of course, carry further and are more resistant to ambient noise, which renders the loud voice especially useful in situations of emergency. Loud speech has greater mean intensity, has a less steeply falling spectral slope, and is pronounced more slowly (see Sluijter & Van Heuven 1996 and references therein). In terms of vocal pitch, we predict that commands should preferably share properties that are held to be characteristic of speaker dominance. Low-pitched voices are strongly associated with dominance, masculinity, authority and competence (e.g. Ohala 1984, Biemans 2000). This is in line with the frequency code (see section 3.3), which says that low pitches are associated with large and dangerous animals across species, and with large and therefore dangerous individuals within the same species, whereas high pitch is associated with small animals and harmless young members within species. It was also found earlier that large pitch movements are associated with strong character, dominance and competence (e.g. Van Bezooijen 1988, Van Heuven & Kirsner 1999). This way of speaking would be the hallmark of leadership (see also Gussenhoven's (2004) effort code – large movements require more effort and therefore a stronger and more powerful individual). Low pitch would show up in acoustic measurements in lowered mean pitch, lower minimum and maximum pitch, while large pitch movements should raise maximum pitch (and therefore also mean pitch) and especially the standard deviation (mean variability) of the pitch as well as the pitch span, i.e., the difference between the highest and the lowest pitch in the utterance. An authoritarian, commanding voice would then be characterised by a lowered low pitches (the L tone segments suggesting a large individual, frequency code) as well as raised high pitches (the H tone segments); the resulting increased size of the pitch changes suggests a strong and energetic individual (effort code).

## 4.2 Method

Four male and four female adult speakers of standard Dutch, aged between 20 and 50, read out sentences as if they were acting in a radio play. The target sentences were direct speech embedded in a preceding matrix sentence that did or did not contain an adverb indicating the tone of voice (or attitude) the speaker had to use in the reporting sentence. The set of sentences was produced once without a specific attitude (neutral), and

second and third time with either a friendly or with a stern (authoritarian) tone of voice. These two attitudes were counterbalanced over the speakers. Participants were instructed to read the sentences as statements (presented with the personal pronoun *ik* 'I') or as imperatives (without the subject pronoun *ik*), again in counterbalanced order. Two sentences were very short, another two were somewhat longer, as shown in Table 4.

**Table 4** Stimulus sentences in experiment 2

*Hij zei* ({*vriendelijk, streng*}): '*ik ga naar huis*'  
 He said ({friendly, sternly}): '(I) go home'

*Hij zei* ({*vriendelijk, streng*}): '*ik lees de krant*'  
 He said ({friendly, sternly}): '(I) read the paper'

*Hij zei* ({*vriendelijk, streng*}): '*ik doe altijd de deur op slot*'  
 He said ({friendly, sternly}): '(I) always lock the door'

*Hij zei* ({*vriendelijk, streng*}): '*ik neem de trein naar Leiden*'  
 He said ({friendly, sternly}): '(I) take the train to Leiden'

Speakers were recorded as described in Haan (2001) and Van Heuven & Haan (2000).

### 4.3 Results

As before, I will analyse the data phonologically (in terms of discrete choices made) as well as phonetically (in terms of gradient differences within the same phonological category). Given Quittart's (1925) observations we expect no systematic differences in choice of tonal categories. Nevertheless, this has to be shown first. Moreover, fine-grained phonetic analyses only make sense if the same phonological categories are compared, which means that a phonological analysis is a necessary prerequisite to the phonetic analysis.

#### 4.3.1 Tone transcription

Using the ToDI conventions (for Tones of Dutch Intonation, see Gussenhoven et al. 2015, Rietveld & Van Heuven 2016: 430-433) a tone transcription was produced of the 384 target utterances by a single transcriber (the present author). Since the target utterances differ in length and number of positions that are potential landing sites for sentence stress, a transcrip-



tion template was adopted that allowed a maximum of four sentence stresses and an initial and final boundary tone (see Table 5). The first potential sentence stress was on the personal pronoun *ik*, which of course remained empty in the case of an imperative. The second landing site would be the verb. Normally, in Dutch (and English) verbs do not receive a sentence stress if they are conjoined with a complement, such as a direct or prepositional object. This tendency is especially strong in the case of verbs with low semantic value, such as those used in our materials (*go, do, take, etc.*). The third location of a sentence stress would be the first noun of the verb complement, which was followed in the two longer structures by a second noun, which was the fourth and last possible location of a sentence stress. The Dutch equivalent for English ‘lock (the door)’ is a complex expression with a semantically bleached verb *doe* ‘do’ plus a prepositional phrase *op slot* ‘on lock’. In combination with the object *deur* ‘door’ the word *slot* is fully predictable after the preposition *op* ‘on’ and will not normally receive a sentence stress, which is why this word is not included in Table 5.<sup>10</sup> The third location for a potential sentence stress only occurs in the longer sentences; this slot is not applicable to the short sentences. The third location is the landing site for a sentence stress on a prosodic dependent, i.e., either the noun *trein*, ‘train’ or the adverb (*altijd* ‘always’). The prosodic heads of the complement phrases in the longer sentences are *Leiden* and *deur* (see below), which are expected to receive sentence stress without a single exception. Sentence stress on a prosodic dependent is optional.

**Table 5** Survey of structural positions that can be realised with a sentence stress in the materials produced in experiment 2

1. subject	2. verb	3. N1	4. N2
(ik)	ga		huis
(ik)	lees		krant
(ik)	neem	trein	Leiden
(ik)	doe	altijd	deur

As a first approximation, I examined the distribution of sentence stresses transcribed in each of the four sentences, separately for statements and imperatives, and broken down further by the three attitudes. The results are shown in Figure 5.

<sup>10</sup> The results contain one exception out of 384 cases, where the speaker stressed the word *slot* (instead of *deur*), which must be interpreted as a contrastive stress.

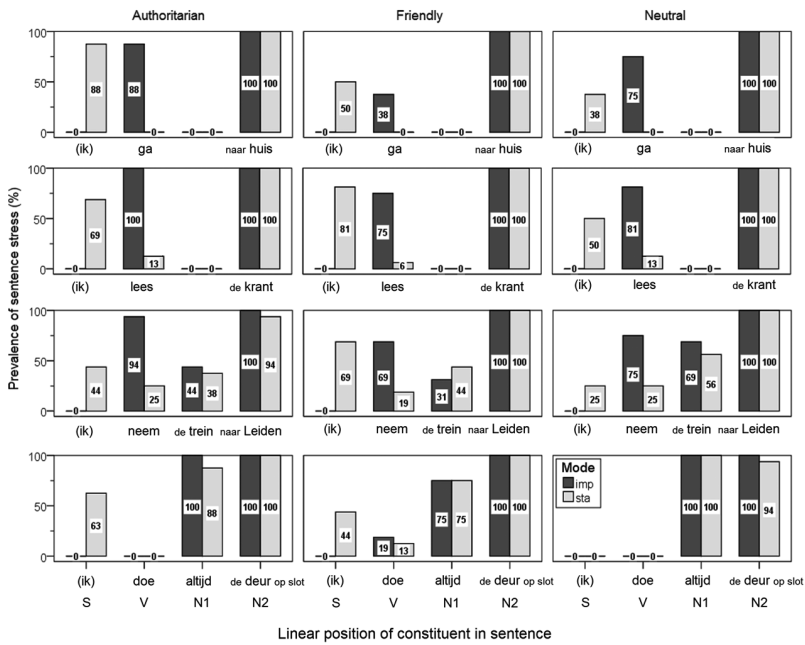


Figure 5 Number of sentence stresses (%) in each of (maximally) four locations in four sentences broken down by attitude and by sentence type (imperative, statement).  $N = 16$  for each bar. No sentence stress is possible on the absent subject (ik) in the imperative mode.

Figure 5 shows that the N2 location (the nuclear position) is nearly always realised with a sentence stress (just two exceptions in 384 cases). In the two short sentences (coded *huis* and *krant*) the location N1 does not exist, which explains why there are no stresses there for these sentences. The two longer sentences, marked *trein* and *deur*, behave rather differently. In the complex NP *de trein naar Leiden* the prosodic head is *Leiden*, while the syntactic head *trein* is the prosodic dependent, which can optionally remain unstressed at the sentence level. This explains why prevalence of stresses on *trein* is 47% as opposed to 99% on the prosodic head *Leiden*. There is no systematic difference in stresses on *trein* between statements (46%) and imperatives (48%), but it would appear that the number of stresses on the dependent is larger in the neutral reading (63%) than in the affect readings (41 and 38% for authoritarian and friendly, respectively).<sup>11</sup> The greatest variability in sen-

11 This effect may be due to the faster speaking rate in the authoritarian and friendly readings than in the neutral reading. We know that optional sentence stresses on prosodic dependents tend to be deleted in faster speech (Caspers & van Heuven 1995).

tence stressing is seen in the verb. The prevalence of stresses on the verb is between 0 and 25% in statements (9% on average) but it increases substantially when the sentence is an imperative, with prevalence between 38 and 100% (59% on average). Here, however, we have to exclude the *deur* sentence. In this sentence the semantically empty verb *doe* is immediately followed by the adverb *altijd*, which attracts the sentence stress: in the imperative it is stressed in 100% of the cases in the authoritarian and neutral reading, and still in 75% of the cases in the friendly reading. The personal pronoun *ik* 'I' should not normally receive a sentence stress. Yet, in the neutral reading, we find the pronoun stressed between 0 and 50% (28% on average) of the cases. In the non-neutral readings, moreover, the number of stressed pronouns is much larger (66 and 61% prevalence for authoritarian and friendly, respectively), which I take to be a reflection of greater personal involvement on the part of the speaker in the affect readings.

As an interim conclusion we may say that there is a statistical tendency for the distribution of sentence stresses to be different between statements and the corresponding imperatives, but the difference is not mandatory. There are still many imperatives in which the speaker does not stress the verb, and there are also tokens of statements in which the speaker stresses the verb where this is not required. Typically, then, there is a tendency towards complimentary distribution of optional sentence stresses on the subject pronoun *ik* and on the verb immediately following it. In the imperative, when the pronoun is absent, the optional sentence stress surfaces on the verb.<sup>12</sup> Interestingly, the tendency to realise optional stresses early on in the sentence is stronger in the affect readings, in which the speaker was instructed to exude authority or friendliness.

In my view, the difference between the statement and imperative stress distribution should not be seen as an inherent property of the sentence type. It falls out as a by-product of the absence of the subject pronoun in the imperative: the speaker feels an urge to realise a sentence stress somewhere in the beginning of the sentence. When there is a subject pronoun, the preferred landing site is the pronoun; when it is not there, the early stress shifts to the next available constituent, i.e., the verb. As a consequence of this we would predict that any verb is more likely to be stressed in sentence-initial position than would be the case in its default V2 position. Inversion of subject and finite in yes-no questions, for instance, should yield an increase in the number of stresses realised on the verb.

12 Yet, in 6 cases (out of 192) stresses were realised on both the pronoun and the verb (5 different types). In all these cases the sentence stress on the verb was a downstepped !H\*(L).

Unfortunately, this hypothesis cannot be checked on Haan's (2001) data. She presents the prevalence of sentence stresses against subject and object nouns in yes/no-questions but does not report stresses on the finite verb. She does mention, however, that 32 out of 199 subject nouns are non-stressed when the verb is preposed, while this occurs only once when the subject is in sentence-initial position (Haan 2001: 108). The preposed finites arguably carry a sentence stress.

Let us now consider the phonological shapes of the melodies realised on the imperatives and statements in the materials. In all, 47 different melodies were found in the materials. Some of these occurred in a relatively large number of cases, others much less, and 17 melodies occurred just once. To reduce variability I treated all melodies that were found fewer than five times in either statements or imperatives, as a rest category. This procedure yielded seven different melodies for imperatives against nine different types for statements. The larger number of statement melodies would be due to the presence of the extra slot for stress on the subject pronoun, which is not available in the imperative.

**Table 6** Phonologically different melodies on imperatives and statements, accumulated over attitudes and lexically different sentences. Melodies with relative frequencies < .03 are lumped together as 'other'.

Imperatives					Statements					
V	N1	N2	N	%	ik	V	N1	N2	N	%
—	—	H*L	28	15	—	—	—	H*L	40	21
H*L	—	H*L	53	28	—	H*L	—	H*L	7	4
					H*L	—	—	H*L	45	23
—	H*L	H*L	22	11	—	—	H*L	H*L	19	10
—	H*	H*L	20	10	—	—	H*	H*L	13	7
H*	—	H*L	19	10	H*	—	—	H*L	15	8
H*L	H*L	H*L	12	6	H*L	—	H*L	H*L	7	4
					H*L	—	H*	H*L	6	3
						—	H*	!H*L	5	3
H*L	—	H*	6	3						
Other	(17 different melodies)		32	17	Other	(20 different melodies)			35	18
Total	(24 different melodies)		192	100	Total	(29 different melodies)			192	100

Note: all initial boundary tones were realised as %L; final boundary tones were always L% but just % (rise nor fall) after H\* on N2.

As is shown in Table 6, the first part of the target utterances (*ik* plus verb) has maximally one sentence stress. In line with the view explained above, let us treat stress on *ik* and on the verb as equivalents, where stress on the verb is the only remaining alternative for an optional stress if the pronoun is absent – as in the imperative. It is obvious, then, that statements and imperatives use very much the same melodies. Even if we allow for small variations in the choice of tone configuration, such as the use of a down-stepped H (indicated by '!H'), we find the same melodies in roughly the same proportions in both sentence modes in over 75% of the materials. In the non-overlapping part of the melodies there are no tone sequences that occur more than once that would be uniquely associated with one mode or the other. Clearly, then, sentence stress placement as well as the choice of tone configuration on the stresses is basically the same for statements and imperatives.

Before moving on to the phonetic implementation of sentence mode, let us briefly consider possible phonological differences between the three attitudes. Table 7 presents the breakdown of melodies (accumulated over the four sentences) by attitude. A larger number of different melodies are found for the affect utterances (35 for authoritarian, 28 for friendly) than for neutral utterances (20). Again, to reduce variability only the more frequent melodic types are listed; melodies with a frequency less than 4 (out of 128 utterances) have been lumped together in a rest category. Boundary tones are not listed in this table as they only varied within the rest category: in the more frequent patterns only %L and L% occur.

This yields five melodies that are shared by all three attitudes in roughly equal proportions. These melodies are comprised mainly of H\*L stresses. One shared melody, however, has a prenuclear H\* stress (N = 33 across attitudes). These shared melodies constitute 62% of the melodic types in the materials.

**Table 7** Distribution of most frequently used melodies broken down by intended attitude. Results have been accumulated over statements and imperatives as well as over four lexically different sentences. Melodies with relative frequencies < .03 are lumped together as ‘other’.

Authoritarian						Friendly						Neutral					
(ik)	V	N1	N2	N	%	(ik)	V	N1	N2	N	%	(ik)	V	N1	N2	N	%
—	—	—	H*L	10	7.8	—	—	—	H*L	29	22.7	—	—	—	H*L	29	22.7
—	H*L	—	H*L	23	17.9	—	H*L	—	H*L	20	15.6	—	H*L	—	H*L	21	16.4
—	—	H*L	H*L	9	7.0	—	—	H*L	H*L	14	10.9	—	—	H*L	H*L	18	14.1
H*L	—	—	H*L	12	9.4	H*L	—	—	H*L	20	15.6	H*L	—	—	H*L	7	5.5
—	—	H*	H*L	12	9.4	—	—	H*	H*L	7	5.5	—	—	H*	H*L	14	10.9
—	H*	—	H*L	13	10.2	—	—	—	—	—	—	—	H*	—	H*L	5	3.9
H*	—	—	H*L	13	10.2	—	—	—	—	—	—	—	—	—	—	—	—
H*L	—	H*	H*L	4	3.1	H*L	—	H*L	H*L	5	3.9	—	H*L	H*L	H*L	11	8.6
H*L	—	—	—	—	—	—	H*L	—	!H*L	4	3.1	—	—	H*	!H*L	4	3.1
Other				32	25.0	Other				29	22.7	Other				19	14.8
Total				128	100.0	Total				128	100.0	Total				128	100.0

A few tentative generalisations may be gleaned from table 7. Firstly, it would appear that authoritarian utterances shun sequences of three H\*L stresses. Secondly, friendly utterances are never realised with a pre-nuclear H\* stress on either the pronoun or the verb. Such melodies, which do occur in neutral and authoritarian utterances, would have a high plateau (called a ‘flat hat’ in earlier grammars of Dutch intonation, ‘t Hart, Collier & Cohen 1990) extending over the entire utterance. The extended flat hat projects an attitude of self-assuredness on the part of the speaker, which may be difficult to reconcile with friendliness.<sup>13</sup>

We may conclude from Table 7 that the choice of tone category does not allow any differentiation between statements and imperatives and is only very weakly associated with differences in attitude.<sup>14</sup> In the remainder of this section we will turn to differences in phonetic implementation as markers of sentence mode and attitude.

13 The high plateau presupposes pre-planning on the part of the speaker. The H\* sentence stress can only be followed by the nuclear (i.e., final and most important) sentence stress. An early H\* communicates to the listener that the speaker knows exactly what he is going to do.

14 A similar conclusion was arrived at by Mozziconacci (1998), who found the H\*L configuration in all emotions and attitudes she studied, and reported weak associations between specific tone configurations and affect (including neutral speech but not authority and friendliness).

#### 4.3.2 *Phonetic analysis of statements and imperatives*

Phonetic properties of the prosody of our materials were measured semi-automatically. As before, utterances were segmented by hand. Fundamental frequency was computed by the autocorrelation method implemented in Praat (Boersma & Weenink 1996), errors were corrected interactively and curves were smoothed with a 10-Hz window. Ten parameters were extracted from the part of the target sentences that began with the verb, i.e., excluding the personal pronoun *ik* in statements:

- Duration of the utterance (in ms)
- Maximum pitch (in ERB)
- Minimum pitch (in ERB)
- Pitch span (maximum minus minimum pitch, in ERB)
- Mean pitch (in ERB)
- Standard deviation of pitch (in ERB)
- Coefficient of variation (CoV) of pitch (standard deviation divided by mean pitch)
- Mean intensity (decibels, dB)
- Standard deviation of intensity (dB)
- Spectral slope (decibels per octave, dB/oct)

In what follows I will present results in two series of analyses. In the first analysis we will take a naïve view and present the effects of attitude and mode on the acoustic parameters across all phonologically different realisations of the target sentences. As a result an imperative with a stressed verb will typically be longer and show more pitch variation than the corresponding statement with the verb unstressed. This type of acoustic analysis examines the effects of phonological differences and of low-level phonetic implementation indiscriminately. In the second analysis we will only include statement-imperative pairs within speakers of utterances that were transcribed with exactly the same sequence of symbols, so that any differences in number, location and shape of sentence stresses and boundary tones are eliminated. This selection narrows the number of pairs down from 192 to 56 and leaves 10 cells out of the total of 48 (8 speakers  $\times$  3 attitudes  $\times$  2 modes) empty. These missing values were replaced by the grand mean for the parameter corrected for the main effects of speaker, attitude and mode. This imputation creates no specific bias and allows us to analyse the effects of attitude and mode within speakers by means of a repeated measures ANOVA. The results of these analyses are found in Table 8.

**Table 8** Summary of RM-ANOVA of all ten acoustic parameters. Factors are Attitude (authoritarian, friendly, neutral) and Mode (statement, imperative). Significant effects are in bold face. P-values are Huyhn-Feldt corrected (not shown in the degrees of freedom).

	Attitude			Mode			Attitude × mode		
	F(2, 14)	p	$\eta^2$	F(1, 7)	p	$\eta^2$	F(2, 14)	p	$\eta^2$
All responses (N = 384)									
Max. pitch	.9	.436	.112	2,8	.140	.382	4.0	.081	.363
Min. pitch	.1	.880	.017	8.8	<b>.021</b>	.556	.1	.947	.008
Pitch span	.7	.532	.086	4.0	.085	.364	1.0	.387	.127
Mean pitch	.0	.950	.005	14.1	<b>.007</b>	.668	.3	.687	.046
St. dev. pitch	.7	.519	.088	3.5	.105	.331	.7	.537	.085
CoV pitch	1.1	.377	.130	1.6	.242	.190	1.3	.349	.140
Mean intensity	2.7	.129	.280	16.2	<b>.005</b>	.698	1.0	.399	.123
St. dev. intensity	1.1	.365	.132	29.0	<b>.001</b>	.805	.0	.996	.001
Slope	.9	.410	.119	2.1	.192	.229	1.5	.258	.176
Duration	3.3	.079	.322	49.9	<b>&lt;.001</b>	.877	1.0	.404	.121
Matched intonations contours (N = 112)									
Max. pitch	1.5	.257	.177	9.2	<b>.019</b>	.567	3.1	.075	.309
Min. pitch	1.6	.243	.183	.5	.490	.070	.2	.747	.032
Pitch span	1.0	.406	.121	.0	.841	.006	1.1	.368	.130
Mean pitch	.7	.510	.092	36.7	<b>.001</b>	.840	.4	.590	.052
St. dev. pitch	1.1	.347	.137	2.7	.144	.279	1.4	.272	.171
CoV pitch	1.1	.363	.133	.0	.880	.003	3.0	.081	.302
Mean intensity	.7	.585	.093	14.8	<b>.006</b>	.679	.4	.664	.053
St. dev. intensity	.7	.452	.095	16.2	<b>.005</b>	.698	1.2	.328	.147
Slope	6.0	<b>.013</b>	463	.6	.460	.080	2.0	.179	.219
Duration	1.6	.234	.188	9.0	<b>.020</b>	.562	.5	.619	.066

Table 8 reveals that none of the acoustic parameters is sensitive to the difference in attitude, with just one exception: there is a modest effect of spectral slope, whereby authoritarian utterances have a more steeply falling (-15.4 dB/oct) slope than neutral utterances (-12.5 dB/oct). This effect runs counter to our expectation and is found only in the subset of matched utterances. Effects of mode are more pervasive. In both the full dataset and in the restricted matched subset we find five parameters that are sensitive to mode. Four out of five are shared between the datasets: mean pitch, mean intensity, standard deviation of intensity, and duration of the utterance. Importantly, the effects of mode on these four shared parameters are of comparable magnitude, which shows that the effect of mode persists even if the possible influence of different location and phonological shapes of sentence stresses and boundary tones is ruled out. Moreover, there is not



a single interaction between mode and attitude that reaches significance, so that we may define fairly simple and straightforward phonetic profiles for the statement versus imperative mode on the basis of four stable parameters with additive main effects only. These are drawn in Figure 6; the additivity of the main effects (i.e. absence of interaction) is seen in the fact that the two lines in each panel basically run parallel courses.

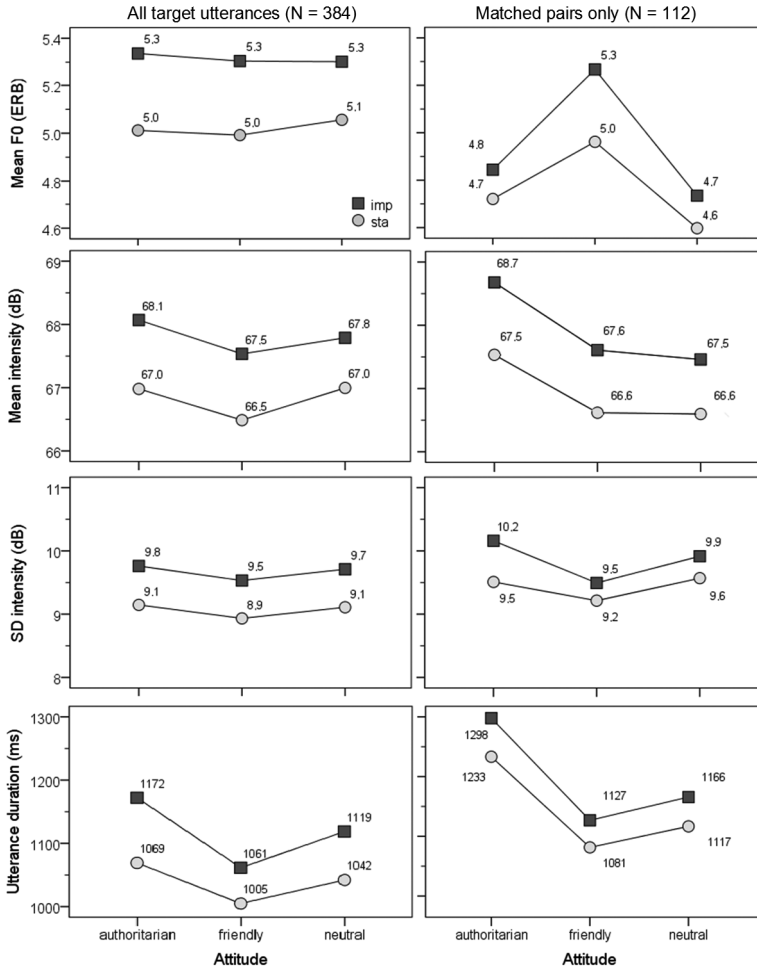


Figure 6 Mean pitch (Fo in ERB), mean intensity (dB), standard deviation of intensity (dB) and duration (ms) as a function of attitude and sentence mode (imperative, statement). Left-hand column presents the data of all 384 target utterances; right-hand column presents the data for only the 62 imperative-statement pairs with matching tone transcriptions.

Imperatives have a mean pitch of roughly .3 ERB higher than statements, the mean intensity is roughly 1 dB stronger. The standard deviation of the intensity is roughly .5 dB larger in imperatives, indicating that there are (slightly) larger differences between loud and less loud syllables. Imperatives have longer durations than the corresponding statements by some 7 per cent, indicating a slightly (but significantly) slower speaking rate in imperatives. The differences (and effect sizes, i.e.,  $p\eta^2$  values) between the imperatives and the overlapping parts of the corresponding statements are not much weaker for the tonally matched sentence pairs than for the total dataset (and in one parameter even larger, i.e., mean pitch). This means that speakers may use alternative strategies to mark imperatives as more forceful speech utterances than statements: they may raise the pitch and loudness and speak more slowly and deliberately in imperatives by inserting more sentence stresses than in statements, or they can just increase the pitch, loudness and duration of the utterance overall, without changing the phonological make-up of the utterance. The predictions made by the frequency and effort codes are not borne out by the effect of sentence mode. The mean pitch of imperatives is higher instead of lower in statements and the pitch span (size of pitch movements) does not differ significantly, even though an increased pitch span was predicted for imperatives. It would appear, therefore, that the overall effect of the imperative mode is just to speak somewhat louder: this raises the overall pitch and intensity and slows down the speaking rate.

It would also appear to be the case that the authoritarian voice is slower and louder than the neutral voice, while a friendly attitude is predominantly conveyed by a higher pitch. These effects, however, failed to reach significance.

## 5 Conclusion

This paper examined the phonological and phonetic effects of three sentence types in Dutch, i.e. statements, (several types of) questions and imperatives. The results have shown that speakers may employ (sequences of) different phonological categories to mark the contrast between the various sentence types, such as the H% high boundary tone at the end of questions or the stressing of the finite verb in the imperative. The use of these phonological markers is optional as long as the sentence type is also cued by lexico-syntactic means. As a result, the association between such phonological means and the sentence type is loose. Moreover, the high

boundary tone is not an exclusive question marker; it also serves to indicate the end of a subclause. The difference between the H% tone in questions and at the end of a subclause is not part of the phonology: this is a continuous rather than a categorical difference. The only difference in the body of the sentence (i.e., the part preceding the final boundary tone) that remains between statements and the various question types is in the details of the phonetic implementation, i.e., the slope of the lower trend lines of the intonation pattern and the relative height of the H\* sentence stress targets.

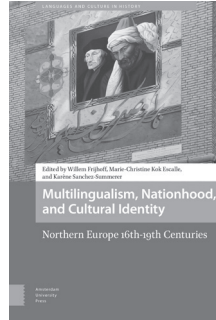
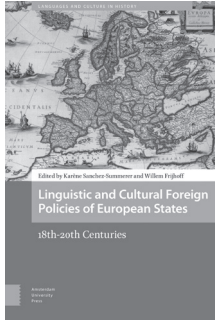
Similarly, the optional stressing of pre-nuclear sentence positions in imperatives is also found in statements (be it less often), and would seem to be a by-product of the deletion of the personal pronoun in the subject position of the imperative. For the statement-imperative contrast, I would argue, once more, that the differences reside in the phonetic implementation, i.e., higher pitch, greater loudness and slower speaking rate (all three suggesting a louder voice), to make imperatives sound more forceful than statements.

### **About the author**

**Vincent J. van Heuven**, University of Pannonia, Veszprém, Hungary  
E-mail: V.J.J.P.van.Heuven@hum.leidenuniv.nl

# Recent verschenen bij uitgeverij AUP

De nieuwe serie **Language & Culture in History** gaat in op de rol van buitenlandse talen bij het ontstaan van het taalkundige en culturele erfgoed van Europa. Er wordt onderzoek gedaan op zowel individueel, regionaal, nationaal en internationaal niveau. De eerste drie titels in de reeks zijn recent verschenen!



*Linguistic and Cultural Foreign Policies of European States. 18th-20th Centuries*, Sanchez-Sommerer en Frijhoff (red.) | hardback | ISBN 978 94 6298 060 0 | € 89.00  
*Multilingualism, Nationhood, and Cultural Identity. Northern Europe, 16th-19th Centuries*, Frijhoff, Kok Escalle en Sanchez-Sommerer (red.) | hardback | ISBN 978 94 6298 061 7 | € 79.00  
*French as Language of Intimacy in the Modern Age*, Van Strien-Charbonneau en Kok Escalle (red.) | hardback | ISBN 978 94 6298 059 4 | € 89.00

Soms ligt het op het puntje van je tong. Vaak weet je dat een bepaalde uitdrukking bestaat, maar kun je er even niet opkomen.

*Met zoveel woorden* van Rik Schutz en Ludo Permentier wijst de weg naar manieren waarop je in het Nederlands iets krachtig en/of beeldend kunt zeggen. Een geweldig bladerboek voor alle taalliefhebbers en onmisbaar voor tekstschrijvers en vertalers.

*Met zoveel woorden*, Rik Schutz en Ludo Permentier | paperback | € 29,95 | ISBN 978 94 6298 180 5 | e-ISBN: 978 90 4853 156 1



Amsterdam  
University  
Press

[www.aup.nl](http://www.aup.nl)

## Functional trade-off of prosody and syntax in question marking?

Vincent J. van Heuven

NT 22 (1): 41–42

DOI: 10.5117/NEDTAA2017.1.HEUE

As a (linguistic) phonetician I am a natural believer in functional explanations of observed regularities. I endorse the view that true explanations should be based on insights that go beyond the boundaries of the specific domain of observation that a theory is intended to account for. It is not enough, for instance, to “explain” phenomena found in one language by showing that the same phenomenon or regularity is shared by more or even all natural languages known at this time. Showing the universality of a regularity found in a specific language is, in this view, not an explanation in its own right but begs the question why many or all human languages possess the same regularity (see, for instance, Lindblom 1979, Boersma 1998). Within this framework of biological functionalism, valid domains of knowledge outside linguistics proper from which explanatory principles can be obtained would be – among many other disciplines – anatomy, neurology, physiology, audiology (studying what the human brain and speech and/or hearing organs can and cannot do), and cognitive sciences (specifically developmental psychology and theories of learning and forgetting).

In her response to my contribution to this special issue of *Nederlandse Taalkunde* Karen De Clercq discusses the observation that the need for prosodic marking of interrogativity is inversely related to the number of lexical and/or syntactic question markers in the sentence. We hypothesized that the need for prosodic question markers would increase monotonically going from statements (ST, none) to wh-questions (WH, question word plus inversion) to yes/no-questions (YN, inversion only) to declarative questions (DE, surface structure identical to ST): DE > YN > WH > ST. We predicted this inverse relationship from a basic functional principle that is found both within and outside language. For instance, plants de-

pend on insects for pollination. Insects are drawn by the flowers that the plant develops. Flowers attract insects by their colour and/or by their smell. It has been known for well over a century (e.g. Boulger 1878) that smell and colour are inversely correlated in (natural, i.e., not cultivated) flowers: white (colourless) flowers are more fragrant than colourful flowers, while monochromous flowers have been found to give off more smell than more variegated flowers – so that in the long run all types of flower tend to be equally attractive to insects (albeit different ones, with different preferences). This principle of functional trade-off is what prompted our hypothesis.

De Clercq points out that the amount and complexity of abstract syntactic machinery that has been postulated in generative theory reflects the same hierarchy that was hypothesized and experimentally confirmed by our study. I am impressed by the ingenuity of the proposals that have been developed by generative syntacticians. Moreover, I am convinced that the proposals were made on autonomous syntactic grounds, and therefore provide independent support for our conclusions with respect to prosodic interrogativity marking in Dutch. I do not rule out the possibility, however, that the complexity of the superstructure (i.e. the multi-layered CP and beyond) needed to differentiate ST, WH, YN and DE hierarchically from one another can in itself be explained in a functional manner. After all, all four structures assume one basic (unmarked) form, i.e., ST. So I would expect that as question types differ more (and at more levels) from the basic form (ST), they require a more elaborate superstructure – which is what De Clercq finds.

## About the author

**Vincent J. van Heuven**, University of Pannonia, Veszprém, Hungary  
E-mail: V.J.J.P.van.Heuven@hum.leidenuniv.nl

## References

- Baker, Mark (1970). Notes on the description of English questions. *Foundations of Language* 6, 197-219.
- Bansal, Rajend K. (1966). *The intelligibility of Indian English*. Ph.D. dissertation, London University.
- Bennis, Hans (1997). Voegwoordvariaties. In: A. van Santen & M. van der Wal (eds.), *Taal in tijd en ruimte*. Leiden: SNL, 353-364.
- Bennis, Hans (2000). On the interpretation of functional categories. In: H. Bennis, M. Everaert & E. Reuland (eds.), *Interface strategies*. Amsterdam: KNAW, 37-53.
- Berg, Rob van den, Carlos Gussenhoven & Toni Rietveld (1992). Downstep in Dutch: implications for a model. In: G.J. Docherty & D.R. Ladd (eds.), *Papers in laboratory phonology II*. Cambridge: Cambridge University Press, 335-359.
- Bezooijen, Renée van (1988). The relative importance of pronunciation, prosody and voice quality for the attribution of social status and personality characteristics. In: R. van Hout & U. Knops (eds.), *Language attitudes in the Dutch area*. Dordrecht: Foris, 85-103.
- Biemans, Monique (2000). *Gender variation in voice quality*. Utrecht: LOT.
- Boersma, Paul (1998). *Functional phonology. Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Boersma, Paul & David Weenink (1996). *Praat, doing phonetics by computer*. Report nr. 136. Institute of Phonetic Sciences, University of Amsterdam.
- Boulger, George S. (1878). Scent and colour in flowers. *Nature* 18, 427-428
- Bresnan, Joan (1972). *Theory of complementation in English syntax*. Ph.D. dissertation, MIT.
- Caspers, Johanneke & Vincent J. van Heuven (1995). Effects of time pressure on the choice of accent-lending and boundary-marking pitch configurations in Dutch. *Proceedings of Euro-speech '95*, 1001-1004.
- Cheng, Lisa (1991). *On the typology of wh-questions*. Ph.D. dissertation, MIT.
- Chomsky, Noam & Howard Lasnik (1977). Filters and control. *Linguistic Inquiry* 8, 425-504.
- Cinque, Guglielmo (1999). *Adverbs and Functional Heads*. Oxford: Oxford University Press.
- Craenenbroeck, Jeroen van (2010). *The syntax of ellipsis. Evidence from Dutch dialects*. Oxford: Oxford University Press.
- Elfner, Emily (2012). *Syntax-prosody interactions in Irish*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Gooskens, Charlotte & Vincent J. van Heuven (1995). Declination in Dutch and Danish. *Proceedings of the 13th International Congress of Phonetic Sciences*, Stockholm, 374-377.
- Gussenhoven, Carlos (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, Carlos, Toni Rietveld, Joop Kerkhoff & Jacques M.B. Terken (2015). *ToDI: Transcription of Dutch intonation*, 2<sup>nd</sup> edition. <<http://todi.science.ru.nl>>
- Haan, Judith (2001). *Speaking of questions. An exploration of Dutch question intonation*. Ph.D. dissertation. Utrecht: LOT.

- Haan, Judith & Vincent J. van Heuven (1999). Male vs. female pitch range in Dutch questions. *Proceedings of the 14th International Congress of Phonetic Sciences*, 1581-1584.
- Haan, Judith & Vincent J. van Heuven (2003). This is a yes/no question? In: P. Fikkert & L. Cornips (eds.), *Linguistics in the Netherlands 2003*. Amsterdam: John Benjamins, 59-70.
- Haan, Judith, Vincent J. van Heuven, Jos J.A. Pacilly & Renée van Bezooijen (1997). On the anatomy of Dutch question intonation. In: J. Coerts & H. de Hoop (eds.), *Linguistics in the Netherlands 1997*. Amsterdam: John Benjamins, 97-108.
- Haegeman, Liliane (1996). Verb second, the split CP and initial null subjects in Early Dutch finite clauses. *Geneva Generative Papers* 4, 133-175.
- Haegeman, Liliane (2012). *Adverbial clauses, main clause phenomena and the composition of the left periphery: the cartography of syntactic structures*. Oxford: Oxford University Press.
- Haegeman, Liliane & Virginia Hill (2013). The syntactization of discourse. In: R. Folli, C. Sevdali & R. Truswell (eds.), *Syntax and its limits*. Oxford: Oxford University Press, 370-390.
- Hart, Johannes 't, René Collier & Antonie Cohen (1990). *A perceptual study of intonation*. Cambridge: Cambridge University Press.
- Heim, Johannes, Hermann Keupdjio, Zoe Wai, Man Lam, Adriana Osa Gómez & Martina Wiltschko (2014). How to do things with particles. In: L. Teddman (ed.), *Proceedings of the 2014 annual conference of the Canadian Linguistic Association*. <<http://cla-acl.ca/actes-2014-proceedings/>>.
- Hermes, Dick J. (1988). Measurement of pitch by subharmonic summation. *Journal of the Acoustical Society of America* 83, 257-264.
- Heuven, Vincent J. van (1994). Introducing prosodic phonetics. In: C. Odé & V.J. van Heuven (eds.), *Phonetic studies of Indonesian prosody*. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, RU Leiden, 1-26.
- Heuven, Vincent J. van (2008). Making sense of strange sounds: (mutual) intelligibility of related language varieties. A review. *International Journal of Humanities and Arts Computing* 2, 39-62.
- Heuven, Vincent J. van (2014a). Acoustic correlates and perceptual cues of word and sentence stress: Mainly English and Dutch. *Proceedings of the 4th International Symposium on Tonal Aspects of Languages*, 211-217.
- Heuven, Vincent J. van (2014b). *Laatste les: Apenkooi*. Universiteit Leiden.
- Heuven, Vincent J. van & Judith Haan (2000). Phonetic correlates of statement versus question intonation in Dutch. In: A. Botinis (ed.), *Intonation: Analysis, modelling and technology*. Dordrecht: Kluwer, 119-144.
- Heuven, Vincent J. van & Judith Haan (2002). Temporal development of interrogativity cues in Dutch. In: Carlos Gussenhoven & Natasha Warner (eds.), *Papers in Laboratory Phonology VII*. Berlin: Mouton de Gruyter, 61-86.
- Heuven, Vincent J. van, Judith Haan & Jos J.A. Pacilly (1998). Global and local characteristics of Dutch questions in play-acted and spontaneous speech. *Proceedings of an ESCA workshop on sound patterns of spontaneous speech*, 139-142.
- Heuven, Vincent J. van & Robert S. Kirsner (1999). Interaction of grammatical form and intonation: Two experiments on Dutch imperatives. In: R. Kager & R. van Bezooijen (eds.), *Linguistics in the Netherlands 1999*. Amsterdam: John Benjamins, 1-96.
- Heuven, Vincent J. van & Robert S. Kirsner (2004). Phonetic or phonological contrasts in Dutch boundary tones? In: L. Cornips & J. Doetjes (eds.), *Linguistics in the Netherlands 2004*. Amsterdam: John Benjamins, 102-113.
- Heuven, Vincent J. van & Agaath M.C. Sluijter (1996). Notes on the phonetics of word prosody. In: R. Goedemans, H. van der Hulst & E. Visch (eds.), *Stress patterns of the world, Part 1: Background*. The Hague: Holland Institute of Generative Linguistics, 233-269.



- Heuven, Vincent J. van & Ellen A. van Zanten (2005). Speech rate as a secondary prosodic characteristic of polarity questions in three languages. *Speech Communication* 47, 87-99.
- Hill, Virginia (2007). Vocatives and the pragmatics-syntax interface. *Lingua* 117, 2077-2105.
- Hill, Virginia (2013). Features and strategies: the internal syntax of vocative phrases. In: Noel, P. & B. Sonnenhauser (eds), *Vocatives! Addressing between system and performance*. Berlin: Mouton de Gruyter, 79-102.
- Hoekstra, Eric & Jan-Wouter Zwart (1994). De structuur van de CP: Functionele projecties voor topics en vraagwoorden in het Nederlands. *Spektator* 23, 191-212.
- Hoekstra, Eric & Jan-Wouter Zwart (1997). Weer functionele projecties. *Nederlandse Taalkunde* 2, 121-132.
- Katz, Jerrold & Paul Postal (1964). *An integrated theory of linguistic descriptions*. Cambridge, Mass.: MIT press.
- Kirchner, Robert S., Vincent J. van Heuven & Johanneke Caspers (1998). From request to command: an experimental study of grammatical form, intonation and pragmatic particle in Dutch imperatives. In: R. Kager & R. van Bezooijen (eds.), *Linguistics in the Netherlands 1998*. Amsterdam: John Benjamins, 135-148.
- Lindblom, Björn E.F. (1980). The goal of phonetics, its unification and application. *Phonetica* 37, 7-26.
- Lindsey, Geoffrey A. (1985). *Intonation and interrogation: Tonal structure and the expression of a pragmatic function in English and other languages*. Ph.D. Dissertation, University of California Los Angeles.
- Mozziconacci, Sylvie J.L. (1998). *Speech variability and emotion. Production and perception*. Ph.D. Dissertation, Technical University Eindhoven.
- Nolan, Francis (2003). Intonational equivalence: an experimental evaluation of pitch scales. *Proceedings of the 15th International Congress of Phonetic Sciences*, 771-774.
- Nooteboom, Sieb (1997). The prosody of speech: melody and rhythm. In: W. Hardcastle & J. Laver (eds.), *The handbook of phonetic sciences*. Oxford: Blackwell, 640-673.
- Ohalá, John J. (1980). On the acoustic origin of the smile. *Journal of the Acoustical Society of America* 68, S33.
- Ohalá, John J. (1984). An ethological perspective on common cross-language utilization of F<sub>0</sub> in voice. *Phonetica* 41, 1-16.
- Qin, Qing (2007). Understanding foreign-accented language. Experiments on the relative contribution of segments and prosody in Dutch-accented Mandarin. M.Phil. thesis, Leiden University Centre for Linguistics.
- Quittart, Leendert J. (1925). *De intonatie van het Nederlands met inbegrip van een vergelijking met de engelse intonatie [The intonation of Dutch including a comparison with English intonation]*. Utrecht: Kemink & zoon.
- Remijns, Bert & Vincent J. van Heuven (2006). Introduction: between stress and tone. *Phonology* 23, 121-123.
- Rietveld, Antonie C.M. & Vincent J. van Heuven (2016). *Algemene fonetiek [General phonetics]*. Fourth revised edition. Bussum: Coutinho.
- Rizzi, Luigi (1997). The fine structure of the left periphery. In: L. Haegeman (ed.), *Elements of grammar: Handbook of generative syntax*. Dordrecht: Kluwer, 281-339.
- Rizzi, Luigi (2001). On the position 'int(errogative)' in the left periphery of the clause. In: G. Cinque & G. Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*. Amsterdam: Elsevier, 267-296.
- Ross, John R. (1970). On declarative sentences. In: R.A. Jacobs & P.S. Rosenbaum (eds.), *Readings in English transformational grammar*. Waltham, Mass.: Ginn, 222-277.

- Selkirk, Elisabeth (2011). The syntax-phonology interface. In: J. Goldsmith, J. Riggle & A. Yu (eds.), *The Handbook of Phonological Theory*, 2nd edition. Oxford: Wiley Blackwell, 435-484.
- Sluijter, Agaath M.C. & Vincent J. van Heuven (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100, 2471-2485.
- Speas, Peggy & Carol Tenny (2003). Configurational properties of point of view roles. In: A-M. Di Sciullo (ed.), *Asymmetry in grammar*. Amsterdam: John Benjamins, 315-344.
- Starke, Michal (2004). On the inexistence of specifiers and the nature of heads. In: A. Belletti (ed.), *Structures and beyond*, volume 3. Oxford: Oxford University Press, 252-268.
- Thorsen, Nina (1980). A study of the perception of sentence intonation – evidence from Danish. *Journal of the Acoustical Society of America* 67, 1014-1030.
- Wang, Hongyan, Ligang Zhu, Xiaotong Li & Vincent J. van Heuven (2011). Relative importance of tone and segments for the intelligibility of Mandarin and Cantonese. *Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong*, 2090-2093.
- Wiltschko, Martina & Johannes Heim (2016). The Syntax of confirmational: a neo performative analysis. In: G. Kaltenböck, E. Keizer & A. Lohmann (eds.), *Outside the Clause. Form and function of extra-clausal constituents*. Amsterdam: John Benjamins. 303-340.
- Wingfield, Arthur (1975). The intonation-syntax interaction: prosodic features in the perceptual processing of sentences. In: A. Cohen & S.G. Nootboom (eds.), *Structure and process in speech perception*. Berlin: Springer, 148-156.