Speech rate as a secondary prosodic characteristic of polarity questions in three languages

Heuven, V.J. van; Zanten-Wervelman, E.A. van

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Speech Communication 47 (2005) 87–99
SPEECH COMMUNICATION
www.elsevier.com/locate/specom

Speech rate as a secondary prosodic characteristic of polarity questions in three languages

Vincent J. van Heuven *, Ellen van Zanten

Universiteit Leiden Centre for Linguistics/Phonetics Laboratory, P.O. Box 9515, Cleveringaplaats 1, 2300 RA Leiden, Netherlands

Received 30 November 2004; received in revised form 22 April 2005; accepted 18 May 2005

Abstract

Questions (almost) universally differ from statements in that the former have some element of high pitch that is absent in the latter. Therefore, the difference in speech melody (intonation) is considered to be the primary prosodic correlate of the contrast. We now pursue the possibility that an other, secondary prosodic correlate may exist that signals the difference between statement and question. We noted in Manado Malay (an Austronesian language) that questions were spoken at a faster rate than the corresponding statements. We then examined speech rate in questions and statements in two Germanic languages, viz. Orkney English and Dutch. In all three languages we find faster speaking rate in questions than in statements, but with different distribution of the phenomenon over the sentence. In Manado Malay, the difference seems restricted to the boundaries of prosodic domains, in Orkney it is evenly spread over the sentence, and in Dutch it is only found in the middle portion of the sentence. Some speculation on possible causes of the rate difference between statements and questions is offered in conclusion.

Keywords: Intonation; Speech rate; Prosodic boundary; Clause type; Statement; Question

1. Introduction

1.1. Question types

It is widely acknowledged that there is not a single human language that does not permit its users to ask questions (e.g. Chisholm, 1982). The speech act of questioning may invoke a variety of lexical, syntactic and prosodic means. By asking an open question, the speaker prompts the addressee to forward a specific piece of information in order to fill in a gap in the speaker's knowledge. Open questions involve the use of question words such as *when*, *what*, and cetera. They are open in the sense that the set of possible answers is virtually unlimited. For example, any name would be an appropriate answer to the open question *What is*

* Corresponding author. Tel.: +31 71 527 2319; fax: +31 71 527 2615.
E-mail address: v.j.j.p.van.heuven@let.leidenuniv.nl (V.J. van Heuven).

0167-6393/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.specom.2005.05.010

your name? Depending on the language, the question type may also involve syntactic changes, such as the inversion of subject and finite verb in the example given.

There is a second main type of questions, called polarity questions, which differs from the open type in the property that it basically allows only two possible responses, positive (*yes*) or negative (*no*). Thus, the polarity question *Is your name John?* is answered by either yes or no. In the present article we will restrict our attention to polarity questions. The reason for this decision is that we will compare statements and questions which do not differ in the number and identity of words. Although subtypes exist, also cross-linguistically, that mark polarity questions by the use of tags or question particles, thereby increasing the number of words in the sentence, polarity question and statement often share exactly the same words, possibly with a change in word order (typically inversion of subject and finite verb) or even without such a change of word order. Since the latter type of question does not differ in any lexico-syntactic way from the statement, is it often called a declarative question. In declarative questions, then, interrogativity is only marked through prosodic means. In Germanic languages such as Dutch and English, there may be a subtle difference in meaning between inversion questions (*Has the last bus left?*) and declarative questions (*The last bus has left?*) in that the latter type suggests that the most likely answer is yes (i.e. speaker asks for confirmation) whilst in the former type the positive and negative possibilities are offered as equally likely (Haan and van Heuven, 2003). In many other languages, however, the inversion type does not occur, so that there the declarative type is the normal way of asking a polarity question.

1.2. Prosody of interrogatives

A considerable body of data has come available on the prosody of interrogative structures. The most noticeable prosodic difference between statements and questions has traditionally been found in the intonation patterns, i.e. the sentence melody. Although languages differ widely in their implementation of interrogativity in terms of sentence melody, it has often been claimed that questions typically involve some element of high pitch which is absent from the corresponding statements (e.g. Bolinger, 1989; Hermann, 1942; Lindsey, 1985). The use of a terminal high boundary tone (*H%*) would probably be the most widespread prosodic question marker across languages. Other forms of high pitch in questions reported in the literature are the suppression of downtrend in Copenhagen Danish (Thorsen, 1978) or the expansion and raising of the peak for a local rise as in many varieties of Southern Italian (Grice et al., 2005). Although prevailing cross-linguistically, the use of high pitch in questions is not universal. Hungarian, for instance, does not use high terminal pitch to signal interrogatives, but differentiates questions from statements in that a terminal rise-fall pitch pattern occurs later, i.e., closer to the end of the utterance, in questions than in statements (Gósy and Terken, 1994). In certain African languages, moreover, the use of pitch in marking interrogativity appears to be reversed. Rialland (2004) presented evidence that in all of the 18 Gur languages in a database of 80 African languages polarity questions are characterized, among other features, by low word tones, falling intonation and lengthening (i.e. a deceleration of speech rate). Rialland's work does not only indicate that the association between pitch and sentence type is to some extent arbitrary, but also shows that the prosodic marking of interrogativity need not be restricted to speech melody; it may well include other prosodic phenomena such as lengthening.

In our own comparative work on word and sentence prosody we invariably found that interrogativity was associated with the use of high pitch, whether locally (high pitch at the end of the sentence, or higher pitch peaks on accents) or globally, i.e. distributed across the entire length of the utterance (van Heuven and Haan, 2000, 2002; van Heuven and van Leyden, 2003). Taking a cue from Rialland's work, we reanalyzed our materials as a first exploration of the possibility that speech rate might be a secondary cue marking interrogativity, next to pitch as the primary

cue. In the present article we will examine three languages for which we have access to statement and polarity question versions of sentences that involve the same words, often in the same order (declarative questions) or with inversion of subject and finite verb (inversion questions). We will study the temporal differences between the clause types (statement, question). The raw materials were collected by students in our own laboratory working on doctoral dissertations on melodic differences between clause types, especially the difference between statements and questions. The first language will be Manado Malay (Stoel, 2005), which is an Austronesian language in which polarity questions are always of the declarative type. In such a language the difference between the two clause types is expressed by prosody only, which led us to expect that prosodic means (both melodic and temporal) will be exploited to the maximum. Second, we will analyze the variety of English that is spoken on the Orkney islands (van Heuven and van Leyden, 2003; van Leyden, 2004). Unlike the Southern British varieties Orkney English marks its accented, focused syllables not with a high-pitched target on the stressed syllable but by a low-pitched target on the stressed syllable followed by high pitch (*H*) on the post-stress syllable. The last language will be Dutch (see also Haan, 2001; van Heuven and Haan, 2000). Dutch is a West Germanic language just like Orkney English but normally marks accents with a high target. In each of these languages we will examine the temporal development of speech rate over the course of the utterance, contrasting statement and question versions of the same lexical sentences. We point out that the materials were collected with the specific view of studying speech melody rather than the effect of sentence type on (local) speech rate. The scope of the materials differs substantially across the three languages, with differences in number of speakers, number of question types, and type of linguistic units that were segmented. Still, we argue that the results show that questions are generally characterized by a somewhat faster speaking rate, whether globally or locally, than the corresponding statements.

2. Questions and statements in Manado Malay

2.1. Introduction

Manado Malay belongs to the Austronesian language family. It is spoken on the isle of Sulawevi (formerly Celebes), which is part of the Republik Indonesia. Typologically, Manado Malay is very different from Germanic languages (and from Indo-European languages in general). As stated above, Manado Malay uses the declarative question type as the general-purpose polarity question.

2.2. Method

Stoel (2005) recorded statements and polarity questions from two young, adult, female Manado-Malay speakers, Lusi and Yessi. Table 1 presents the materials.

The sentences were presented to the speakers on a sheet of paper, in conventional Manado-Malay spelling (see Table 1). Speakers produced the tokens in isolation, using a Shure SM10A head-mounted close-talking microphone and a Sony WM-D6C analog cassette recorder. During the recordings the speaker and the experimenter were seated in a quiet room. Speakers were instructed to pause and inhale before uttering the next

Table 1
Manado-Malay sentences

(a)	Tu Nana pe anak so basar [tu 'nana pe 'anak so ba'sar] the Nana poss. child part. big 'Nana's child is already big'
(b)	Ngana pe tamang so makang ['ŋana pe 'tamaŋ so 'makaŋ] You poss. friend part. eat 'Your friend has already eaten'
(c)	Dia bacirita capat ['dia baci'rita ca'pat] pro-3sg speak fast 'he/she speaks fast'
(d)	Tu surat dia so balas [tu 'surat 'dia so 'balas] the letter pro-3sg part. answer 'The letter has been answered by him/her'

sentence. One speaker recorded the eight sentences twice, the other just once.

The recordings were transferred to computer disk (16 kHz, 16 bit) and segmented into syllable-sized units by hand with the aid of Praat speech processing software (Boersma and Weenink, 1996). Austronesian languages have a simple CV(C) syllable type and avoid consonant clusters, which renders syllabification a rather easy task. Conventional segmentation criteria were used as laid down in the protocol established for the segmentation of Dutch utterances by van Zanten et al. (1991, 1993).

2.3. Results

Fig. 1 presents the overall speech rate (in syllables per second, syll/s) in the utterances for each of the two speakers separately broken down by clause type (question, statement).

The results indicate that the questions were faster overall in the recordings from both speakers (prima facie, the difference is larger for speaker Lusi than for speaker Yessi). The effect of clause type is significant by a two-way ANOVA with speaker and clause type as factors, $F(1,20) = 8.1$ ($p = .010$). However, the effect of speaker is insignificant, as is the speaker × clause-type interaction ($F(1,20) < 1$ for both terms).

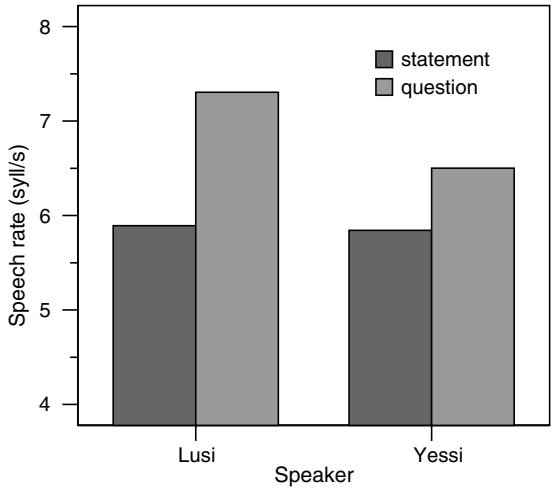

Fig. 1. Speaking rate (syll/s) in statements and declarative questions for two speakers of Manado Malay.

In order to determine the exact location of the temporal difference between statements and questions, Fig. 2 presents the syllable durations for each of the four lexically different sentence types with separate lines for the statements and the (declarative) questions.

Fig. 2 shows that, across the entire duration of the utterances, there is a tendency for syllable durations to be shorter in questions than in the corresponding statements. The difference is very small, except in the words preceding a prosodic boundary. The faster rate is clearest in the words immediately preceding the sentence boundary. The difference is largest in the last syllable before the boundary; a smaller but still substantial difference is found in the penultimate syllable but only if this syllable is stressed (as is the case in *makang* and *balas* but not in *basar* and *capat*). It would appear, then, that the domain of final lengthening here is the pre-boundary foot, which prosodic unit begins with the last stressed syllable before the boundary and is optionally followed (but not preceded) by an unstressed syllable.

There is also pre-boundary lengthening, with a difference between statements (stronger lengthening) and questions (weaker lengthening), for sentence-internal prosodic boundaries (i.e. I-domain boundaries) as is seen after *anak* in sentence (a). It appears that sentence-internal pre-boundary lengthening only takes place if the boundary is in the second half of the prosodic utterance, i.e. in sentences (a) and (b) only but not in (c) and (d).

2.4. Conclusion

The empirical basis of this study is very limited, with only 24 tokens in all collected for only two speakers. Still, the results allow a straightforward conclusion to be drawn. The (declarative) question type in Manado Malay is spoken faster than the corresponding statement. Moreover, the faster rate is not uniformly distributed over the course of the utterance but seems limited to lower-order prosodic domains immediately preceding a deeper prosodic boundary, i.e. the last foot of an intonation domain or utterance.

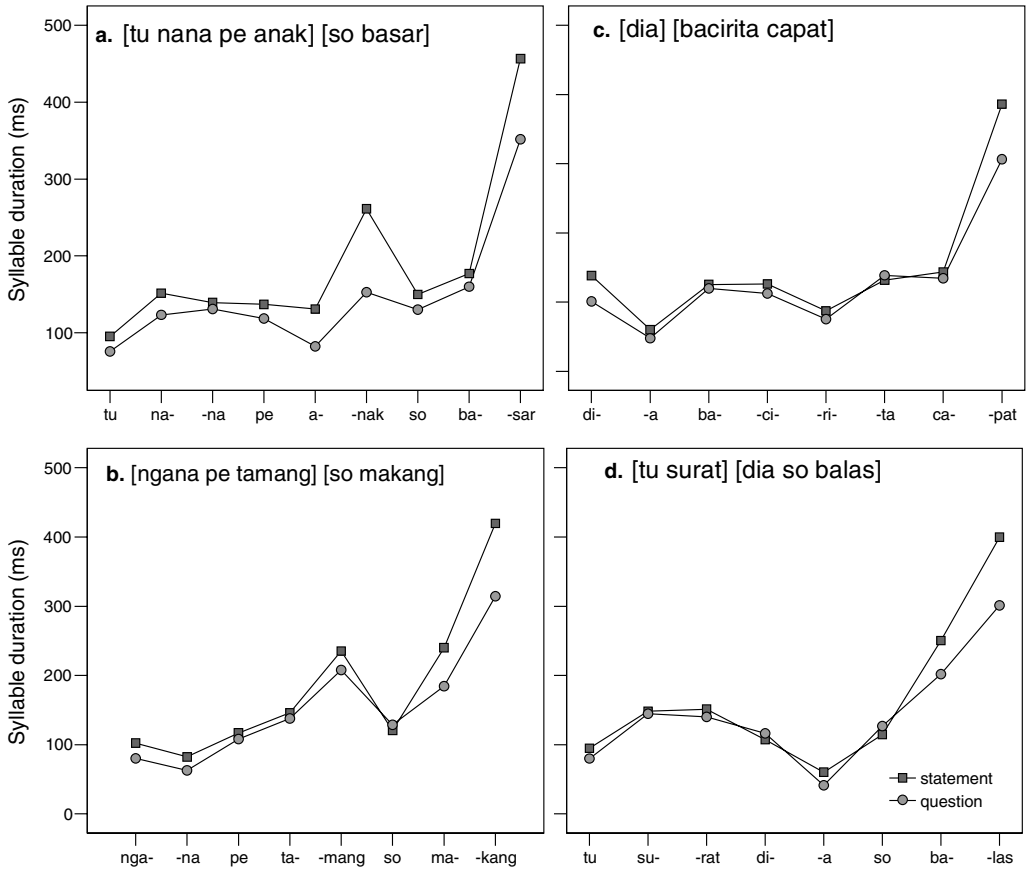

Fig. 2. Syllable duration (ms) of four Manado-Malay sentences broken down by clause type (question/circles versus statement/squares). Utterances were recorded from two native speakers. Square brackets demarcate I-boundaries.

In the next section we will examine the difference between statements and questions in Orkney English. For this language we exploit a larger database of utterances, distributed over a much larger set of speakers.

3. Questions and statements in Orkney dialect

3.1. Introduction

Orkney dialect is a variety of English spoken on the isles of Orkney, off the North-East coast of Scotland. Prosodically it is different from Southern British English in that Orkney dialect marks its accents by a low pitch on the accented syllables followed by a pitch rise on the next unaccented syllable. Standard English marks accented words by realizing a high target on the stressed syllable. It is not clear at this stage whether the difference between Orkney dialect and Standard English is phonological (L^* versus H^*) or whether there is basically one and the same target configuration which is merely time-shifted so as to be aligned differently with the segments.

3.2. Method

In the course of her dissertation work, van Leyden (2004) collected tokens of statements and inversion questions from eleven male and seven female native speakers of Orkney English, aged

Table 2
Statements and questions collected for Orkney English

Statement	Inversion question
1. There are many gardens in Bergen	Are there many gardens in Bergen?
2. There are many houses in Bergen	Are there many houses in Bergen?

between 30 and 50. Each speaker recorded two statements and two corresponding questions, as indicated in Table 2.

The four target sentences were recorded together with 14 other sentences. Stimuli had been printed individually on cue cards, and were presented to each speaker in a different random order. Speakers were recorded onto a minidisk (Sony MZ-R35) in their own home or work place, using a head-mounted Shure SM10A close-talking microphone. Each speaker recorded the list of materials twice; only the second tokens were analyzed, unless an individual item contained a disfluency, in which case the first-recorded token was substituted instead.

The materials were played back, analog-to-digital converted (16 kHz, 16 bit) and analyzed using Praat software.

3.3. Results

Typically the vowels in *There are* are reduced in the statement but tend to keep their full quality (and duration) when the word order is inverted. Therefore we summed the durations of the three accented words (*many*, {*houses, gardens*}, *Bergen*) in the target utterances for each of the 18 speakers separately for questions and statements, and then computed speaking rate in syllables per second. Fig. 3 plots the speech rate in the statements horizontally against the rate in the corresponding questions vertically, for each of the 18 speakers. The diagonal in the graph represents the situation in which the speech rate of statements and questions is equal. The figure demonstrates that the questions are generally faster (above the diagonal) than the corresponding statements. Only four out of 18 individuals have slower questions, but for these four the difference is very small. When the question is faster, the difference tends to assume much larger values. Interestingly, three of the four speakers who have slower questions than statements are male (squares); only one is female (circles). The effect of clause type (statement versus question) is highly significant but we will defer statistical analyses until we have broken the results down to the level of individual syllables.

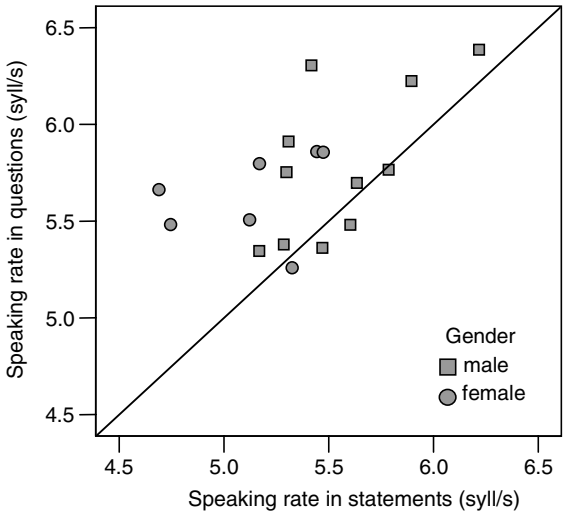

Fig. 3. Speaking rate in accented words (syll/s) in statements (horizontal) plotted against rate in inversion questions (vertical) produced by 11 male (squares) and 7 female (circles) speakers of Orkney dialect.

A more detailed breakdown of the effects of sentence type on the durations of statements and questions is given in Fig. 4, which presents the durations of the successive syllables, as we did for Manado Malay. Separate panels are provided for male and female speakers. It was virtually impossible to segment the words *there* and *are* in the statement version of the sentences; for this reason we have included the onset portion of the statements and questions as one fragment {*there, are*}.

Fig. 4 shows that, indeed, the onset portion of the sentence is the only apparent exception to the general finding that syllable duration is shorter (i.e. speech rate is faster) in questions than in statements for male speakers, and even more clearly for female speakers. The onset of the question (*Are there*) typically contains two full syllables, whilst

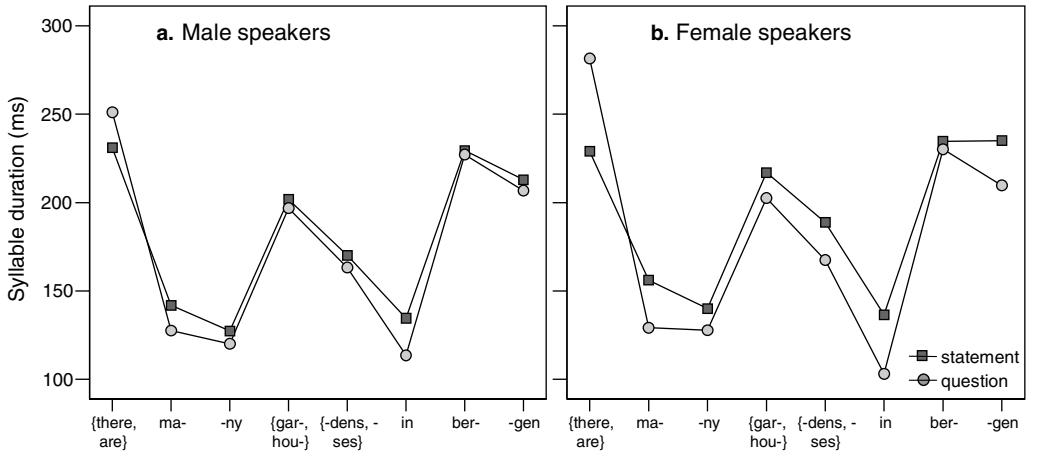

Fig. 4. Syllable duration (ms) in Orkney English statements (squares) and inversion questions (circles) spoken by male (panel a) and female (panel b) speakers.

the beginning of the statement is reduced to [ðərə], with the full vowels replaced by schwas. Therefore only the summed duration of the three accented words (as in Fig. 3) was submitted to a repeated-measures RM-ANOVA with syllable position, lexical type (*houses*, *gardens*) and clause type (statement, question) as within-subject factors. The effect of syllable position was highly significant, $F(3.1, 51.9) = 132.6$ ($p < .001$, Huynh–Feldt corrected). The durations differ significantly for the *houses* and *gardens* sentences, $F(1, 17) = 20.0$ ($p < .001$), but interactions involving this lexical variation were not significant (for this reason we did not break down the results by this factor in Figs. 3 and 4). Crucially, the effect of clause type also reaches significance, $F(1, 17) = 15.9$ ($p = .001$). However, the two-way interaction between clause type and syllable position is not significant, $F(5, 85) = 1.5$ (ins.), indicating that the difference in, for instance, the sentence-final word *Bergen* is not larger (or smaller) than that in other target words.

The effect of clause type on the syllable and word duration in Fig. 4 seems to be larger, prima facie, for the female speakers than for the males. The effect of gender cannot be tested by the RM-ANOVA since this technique requires a fully orthogonal design. In our materials the number of male and female speakers is not the same, so that we had to take recourse here to a simpler type of analysis, viz. a independent-samples *t*-test on the difference in speech rate (as in Fig. 3) between the statements and the corresponding questions per speaker. The analysis shows that the effect of gender is significant but only if we accept one-tailed testing, on the grounds that women are expected to mark the difference between the clause types more saliently than men (following Haan, 2001; Haan and van Heuven, 1999; see discussion section), $t(16) = 2.0$ ($p = .032$, one-tailed).

3.4. Conclusions

As was the case for our two Manado-Malay speakers, we have now determined a similar effect of clause type on speech rate in a completely different language, i.e. the Orkney dialect of English. In Orkney English, too, questions are spoken faster on average than the corresponding statements. The effect proved fairly robust across the set of 18 speakers, with female speakers marking the temporal contrast more clearly than men.

4. Dutch

The largest dataset at our disposal that bears on the issue of prosodic correlates of statements and questions, is for Dutch. This is a set of 600 sentences spoken in isolation or in short paragraphs,

with a complete tone transcription, tonal segmentation, and stylized pitch contours.

4.1. Method

The materials comprised the statement versions and two different types of polarity questions, i.e. inversion questions and declarative questions, of two lexically different sentence types. Speakers were instructed to read out the materials as if they were actors in a radio play. The sentences were either the first or last of a two-sentence paragraph or occurred in isolation. The set of stimuli is given in Table 3.

Recordings were made in a sound-proofed studio using a Sennheiser MKH 416 condenser microphone onto digital audio tape (48.1 kHz). Five male and five female speakers produced the materials twice, yielding a data set of 200 tokens of statements and 200 tokens of each of the two questions types. The recordings were downsampled to 16 kHz, excerpted from their spoken context and acoustically analyzed.

Table 3
Stimulus materials for Dutch

Clause type	Length	
Statement	Short	Renée heeft nog vlees over [rə'ne heft nɔχ 'fles 'ovər] 'Renee has still meat left'
	Long	Marina wil haar mandoline verkopen [ma'rina ʋɪl har mɑndo'linə vər'kopə] 'Marina wants her mandolin sell'
Inversion Q	Short	Heeft Renée nog wat vlees over? [heft rə'ne nɔχ ʋɑt 'fles 'ovər] 'Has Renée still any meat left?'
	Long	Wil Marina haar mandoline verkopen? [ʋɪl ma'rina har mɑndo'linə vər'kopə] 'Wants Marina her mandolin sell?'
Declarative Q	Short	Renée heeft nog vlees over? [rə'ne heft nɔχ 'fles 'ovər] 'Renee has still meat left?'
	Long	Marina wil haar mandoline verkopen? [ma'rina ʋɪl har mɑndo'linə vər'kopə] 'Marina wants her mandolin sell?'

Fundamental frequency (F_0) traces were computed (autocorrelation method) using the Praat speech processing software. The F_0 curves were stylized by hand by locating the F_0 minima and maxima of rise-fall pitch movements on the stressed syllables of the subject and the object NP. The onset and offset frequencies of the terminal boundary tone (L% or H%) were determined, as well as the pitch at the onset of the sentence. The time-frequency coordinates of the pivot points in the stylized pitch contours were stored in a database. Also, the location of the beginning and ending of the onset and vocalic nucleus in the stressed syllables of the (potentially) accented words were stored, as were the segmentation points for all the syllables near the end of the target utterance, i.e. all the syllables after the last potentially accented word.

4.2. Results

As a first approximation, we present the speaking rate in the sentence broken down by sentence type (statement, inversion question, declarative question) and broken down further by the two lexically different contents of the target sentences (*Renée* sentence, with relatively short words; *Marina* sentence, with longer words, see Table 3). In Fig. 5a we plot rate in statements (horizontal) against rate in the inversion questions (vertically) for the *Renée* and *Marina* sentences as produced by each of the 10 speakers. No breakdown by speaker gender is attempted as this factor proved insignificant (see below). Again, points above the diagonal represent faster questions than statements. The analogous Fig. 5b presents the scatterplot for the rate in statements against the rate in declarative questions.

Fig. 5 overall shows that the *Marina* sentences are faster than the *Renée* sentences. This is only to be expected since longer words, containing more unstressed syllables, tend to be spoken faster than short words (Nooteboom, 1985). The figure also shows that the question sentences tend to be spoken slightly faster than the corresponding statements. The effect is clearer in the longer *Marina*

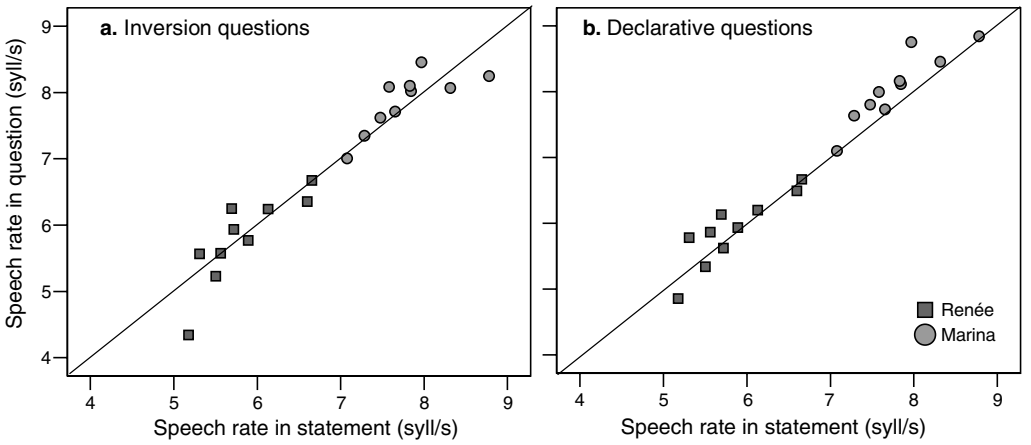

Fig. 5. Overall speaking rate in Dutch statements (syll/s) plotted against rate in corresponding inversion questions (panel a) or in declarative questions (panel b) broken down by short (*Renée*) versus longer (*Marina*) sentences for each of 10 speakers.

sentences than in the *Renée* sentences; also, the faster rate is more apparent in the declarative questions (panel b) than in the inversion questions (panel a). An RM-ANOVA on speech rate with sentence type (statement, inversion question, declarative question) and length (*Renée*, *Marina*) as within-subjects factors, and gender of speaker as a between-subjects factor, indicated no effect of gender ($F(1,8) < 1$), a very strong effect of length, $F(1, 13.2) = 704.8$ ($p < .001$, Huyhn–Feldt corrected), and a smaller effect of sentence type, $F(2, 13.2) = 4.8$ ($p = .025$, Huyhn–Feldt corrected). There were no significant interactions. Paired *t*-tests show that the declarative questions are significantly faster than the statements, $t(9) = 2.8$ ($p = .011$, one-tailed); the difference between inversion questions and statements fails to reach significance, $t(9) = .3$ (ins.). The faster rate in declarative questions is significant in the longer *Marina* sentences, $t(9) = 3.9$ ($p = .002$, one-tailed), but not for the *Renée* sentences, $t(9) = .3$ (ins.). In the *Marina* sentences the speaking rates are 8.1 versus 7.8 syll/s in statements and declarative questions, respectively; in the *Renée* sentences these rates are 5.9 versus 5.8 syll/s.

In order to locate the parts of the target utterances where the faster rate was achieved in the question sentences, Fig. 6 plots the mean syllable duration of successive parts of the utterances with separate lines for the statement and the two question types.

The database does not provide a complete segmentation of the utterances into syllables. The first fragment contains the one or two syllables preceding the stressed (capitalized) syllable of the subject (*maRIna*, *reNÉE*); when two syllables are included, as in the inversion questions (*heeft re*... and *wil ma*...), the mean syllable duration has been plotted. The stresses in *maRIna* and *reNÉE* are separate syllables as are those in *mandoLIne* and *VLEES*, as well as each of post-nuclear syllables following these stresses. The stretch between the stresses in the subject and object has not been segmented into separate syllables. For this portion of the utterance the mean syllable duration has been plotted instead.

Fig. 6 shows that all portions of the sentence have virtually the same duration for statements and declarative questions, with just one exception. The only discrepancy is found in the stretch between the stressed syllable on the subject and that on the object, i.e. the portion *-na wil haar mando-*, which is some 70 ms (5×13 ms) shorter when occurring in the declarative question. Paired *t*-tests show that only this difference is significant, $t(9) = 7.3$ ($p < .001$). The effect is so robust for the *Marina* sentences that it is found for each of the 10 speakers. The figure also shows that the inversion questions (dotted lines) are hard to compare with the statements, probably due to different numbers of syllables, and differences in segment sequences.

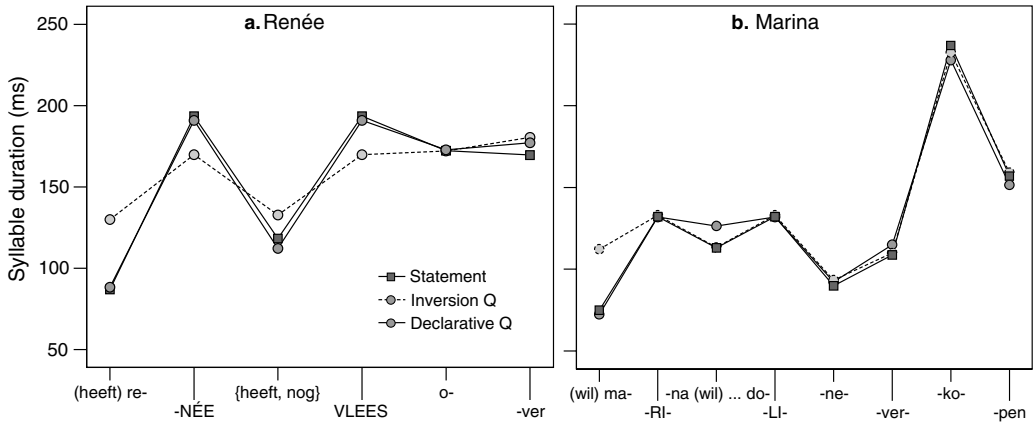

Fig. 6. Duration (ms) of individual successive syllables in *Renée* and *Marina* sentences, broken down by sentence type (statement, inversion question, declarative question). The initial fragments *re-* and *ma-* contain two syllables (*heeft re-* and *wil ma-*, respectively) in inversion questions. The fragment *-na…do–* contains four (inversion question) or five (other types) unstressed syllables: *-na (wil) haar man-do-*.

Speech rate may be affected by accentuation. When a word is accented in Dutch, all its syllables, whether stressed or unstressed, are proportionally lengthened (Eefting and Nooteboom, 1991; Sluijter and van Heuven, 1995, 1996). From this, we would expect the middle portion of the sentence to shrink in duration when the subject or the object is deaccented. Using the expert transcriptions provided in the database (Haan, 2001) we established that only three out of the 200 Marina sentences did not contain accents on both subject (*Marina*) and object (*mandoline*). Of these, the subject was deaccented once in statements, and twice in declarative questions. The effect of sentence type on the duration of the middle fragment of the sentence remains unaffected when we reran the paired *t*-test on the remaining 197 cases, $t(9) = 6.8$ ($p < .001$).

4.3. Conclusions

Secondary markers of interrogativity are subtler in Dutch than in the other languages discussed earlier. Nevertheless, we found faster speaking rate for questions (especially of the declarative type) in the middle portion of the sentence. The effect persisted, even if effects of deaccentuation of either subject or object were kept constant. Other than in Orkney English, there were no indications that women mark the contrast between the clause types more clearly than men.

5. Overall conclusions and general discussion

We have looked for a secondary phonetic correlate of interrogativity in three languages, viz. Standard Dutch, Orkney English dialect, and Manado Malay. The secondary correlate of interrogativity examined in this paper is duration (or speech rate). We will briefly summarize our main findings and then speculate on possible mechanisms underlying the effects that we found.

In each of the three languages studied questions were characterized by some local or global acceleration relative to the corresponding statements. In Manado Malay, the acceleration in questions was concentrated in pre-boundary words; the effect was relatively large, amounting to some 15% difference overall, but yielding much higher values in the restricted domains. In Orkney English the faster rate seems evenly spread over the entire utterance; here the global effect was smaller than in Manado Malay, i.e. some 5% difference in duration. In Dutch acceleration was found only in the middle portion of the utterance between the accents on the subject and on the object of the sentence. This is a local rather than a global effect, and it is small, i.e. some 10% difference locally in the *Marina* sentences, and less than that in the *Renée* sentences. Further analysis showed that the duration effect in Dutch was not confounded by differences in accent structure between statements and questions.

The effects of clause type on speech rate seem robust enough in Manado Malay and Orkney English; in Dutch they are much weaker. Nevertheless, we would like to argue that the effect also exists in Dutch. If the effect is to be found at all, it should show up in the declarative question type, for two reasons. First, the sequence of words and segments is identical in statements and declarative questions, so that no confounds can arise due to inversion of words. Second, the difference between statement and declarative question can only be expressed through prosodic means, so that the need for a secondary acoustic cue to interrogativity would make sense. Although the effect of interrogativity on local speech rate was insignificant in the shorter *Renée* sentences, it was in the same direction as in the *Marina* sentences, and the effect remains significant when the two lexically different sentence types are collapsed. Of course, the findings of the present pilot study need to be confirmed by temporal analysis of questions and statements in many other language samples. On the strength of the results obtained so far we expect that the preferred pattern cross-linguistically will be faster speaking rate for polarity questions.

Let us now speculate briefly on possible reasons why questions might be signaled by faster speaking rate than statements. As mentioned in the introduction, there is a clear bias in the languages of the world to use high pitch as a correlate of questions and low pitch in statements. In the three languages targeted in the present paper this pitch difference was systematically found. Stoel (2005, pp. 101–119) shows that polarity questions in Manado Malay invariably end in a high terminal boundary *H*%, which is absent in statements and open questions. In Orkney English the mean pitch of the polarity questions is appreciably higher than that of the corresponding statements, even if the final boundary tones are discarded from the pitch analysis (van Leyden, 2004, p. 65; van Heuven and van Leyden, 2003). In the Dutch materials the declarative questions ended in *H*% in 100% of the cases and inversion questions in 96%. The pitch of the part of the Dutch utterances before the final boundary tone (*H*% or *L*%) was higher in questions than in statements, especially in the declarative questions, in terms of onset pitch, minimum and maximum pitch and mean pitch (Haan, 2001; van Heuven and Haan, 2000).

The use of high pitch in questions has been explained ethologically by Ohala (1984), as a phonologized remnant of animal behavior. When asking a question, the speaker should be polite or subservient to the addressee. Small (harmless) creatures have higher pitches, and make faster movements, than large (dangerous) creatures. Therefore, by raising his pitch, the speaker pretends to be smaller than he physically is, which is then taken as a sign of submission or politeness to the addressee. By the same token, faster movements, i.e. fast speech rate, would convey the same message from speaker to hearer. This would seem to tie high pitch and fast rate together in the signaling of questions. The same positive correlation between speech rate and pitch has been found in speech perception research. Rietveld and Gussenhoven (1987) asked listeners to judge the speech rate of original and pitch-manipulated versions of sentences, and observed that higher-pitched versions of temporally unchanged utterances were judged to be faster.

An alternative account of the association of high pitch and fast rate with questions may be derived from (Bolinger, 1964, 1989), who claims that statements and questions are characterized universally by a dichotomy between relaxation (low, falling pitch) and tension (high, rising pitch), respectively. It would seem a natural extension of this principle to claim that slow rate or deceleration would then be another correlate of relaxation, and that high rate and acceleration go together with tension. Relaxation at the end of a statement indicates that the speaker has finished his utterance. Ending an utterance with tension signals that the speaker has not finished yet, and that the utterance so far is only the first part of a longer structure. It has in fact been suggested in the older literature (Kretschmer, 1938) that polarity questions could be conceived of as the first part of a compound sentence involving both the suggestion of a proposition followed by its logical complement. Thus, the question *Is your name John?* would be the first part of a longer compound question *Is you name John, or is it not?* This would explain why polarity questions end in a high tone; the high boundary tone would not so much be a

marker of interrogativity but of continuation. It would also explain why polarity questions but not open questions (nor compound questions) end on a high tone. And ultimately, it would explain why speech rate in the polarity question would be faster. Longer sentences generally have higher speech rates than shorter sentences, just as words are spoken faster as they are longer (Lindblom et al., 1981; Grosjean, 1983; Nooteboom, 1985). Although the overall effect of faster speech rate in polarity questions would be explained by this view, certain local effects would still go unaccounted for. For instance, Lindblom et al. (1981) would predict that the faster rate would be noticeable especially in the beginning of the sentences; it would not predict acceleration in the middle portion of the sentence.

We reiterate that the association between high pitch and fast rate with questions is not universal (see work by Rialland discussed in the introduction). Even though linking high (and fast) with interrogativity may be the preferred association in the languages of the world (Gussenhoven, 2004), the fact that counter examples exist, shows that the association is ultimately arbitrary and conventional. But even in the exceptional languages cited by Rialland, the coupling of high pitch and fast rate, and of low pitch and slow rate, is maintained.

Finally, there is the issue whether women are better communicators than men. In Manado Malay only female speakers were recorded, so that this dataset does not speak to the issue. In Orkney English, the women displayed a significantly larger effect of clause type on speech rate than the men. In Dutch, however, no such effect was apparent. The present study, therefore, does not provide a coherent set of results, so that we have to relegate this issue to future research.

Acknowledgements

We are grateful to Ruben Stoel, Klaske van Leyden and Judith Haan for making parts of their materials available for the purpose of this study, and to two anonymous reviewers for their very helpful criticisms. We also thank Jos Pacilly of the ULCL Phonetics Laboratory for his assistance in the semi-automatic extraction of acoustical data from the recorded materials.

References

Boersma, P., Weenink, D., 1996. Praat, doing phonetics by computer. Report nr. 132. Institute of Phonetic Sciences, University of Amsterdam.

Bolinger, D.W., 1964. Intonation as a universal. In: Proc. 9th Internat. Congress of Linguists, Cambridge, Mass., Mouton, The Hague, 1962, pp. 833–848.

Bolinger, D.W., 1989. Intonation and its Uses. Stanford University Press, Palo Alto, CA.

Chisholm, W.S., Jr. (Ed.), 1982. Interrogativity. A Colloquium on the Grammar, Typology and Pragmatics of Questions in Seven Diverse Languages. John Benjamins, Amsterdam.

Eefting, W.F.Z., Nooteboom, S.G., 1991. The effect of information value and accentuation on the duration of Dutch words. J. Acoust. Soc. Amer. 89, 412–414.

Gósy, M., Terken, J., 1994. Question marking in Hungarian: timing and height of pitch peaks. J. Phonet. 22, 269–281.

Grice, M., D'Imperio, M., Savino, M., Avezani, C., 2005. Strategies for intonation labeling across varieties of Italian. In: Jun, S.-A. (Ed.), Prosodic Typology: The Phonology of Intonation and Phrasing. Oxford University Press, Oxford, pp. 261–389.

Grosjean, F., 1983. How long is the sentence? Prediction and prosody in the on-line processing of language. Linguistics 21, 501–521.

Gussenhoven, C., 2004. The Phonology of Tone and Intonation. Cambridge University Press, Cambridge.

Haan, J., 2001. Speaking of questions. An exploration of Dutch question intonation. LOT dissertation series 52, LOT, Utrecht.

Haan, J., van Heuven, V.J., 1999. Male vs. female pitch range in Dutch questions. In: Proc. 14th Internat. Congress of Phonetic Sciences, San Francisco, pp. 1581–1584.

Haan, J., van Heuven, V.J., 2003. This is a yes/no question? In: Fikkert, P., Cornips, L. (Eds.), Linguistics in the Netherlands 2003. John Benjamins, Amsterdam, pp. 59–70.

Hermann, E., 1942. Probleme der Frage, 2 Vols., Vandenhoeck & Ruprecht, Göttingen.

Kretschmer, P., 1938. Der Ursprung des Fragetons & Fragesatzes. In: Ballini, A. et al. (Eds.), Scritti in onore di Alfredo Trombetti, Ulrico Hoepli, Milano, pp. 27–50.

Lindblom, B., Lyberg, B., Holmgren, K., 1981. Durational Patterns of Swedish Phonology: Do They Reflect Short-Term Motor Memory Processes. Indiana University Linguistics Club, Bloomington, IN.

Lindsey, G.A., 1985. Intonation and interrogation: tonal structure and the expression of a pragmatic function in English and other languages. Ph.D. Thesis, University of California, Los Angeles.

Nooteboom, S.G., 1985. A functional view of prosodic timing in speech. In: Michon, J.A., Jackson, J.L. (Eds.), Time, Mind, and Behavior. Springer, Berlin, pp. 242–252.

Ohala, J., 1984. An ethological perspective on common cross-language utilization of F0 of voice. Phonetica 41, 1–16.

Rialland, A., 2004. A typology of question prosody in African Languages. Lecture presented at the Conference on Tone and Intonation in Europe TIE, Santorini, Greece, 9–11 September 2004.

Rietveld, A.C.M., Gussenhoven, C., 1987. Perceived speech rate and intonation. J. Phonet. 15, 273–285.

Sluijter, A.C.M., van Heuven, V.J., 1995. Effects of focus distribution, pitch accent and lexical stress on the temporal organisation of syllables in Dutch. Phonetica 52, 71–89.

Sluijter, A.C.M., van Heuven, V.J., 1996. Spectral balance as an acoustic correlate of linguistic stress. J. Acoust. Soc. Amer. 100, 2471–2485.

Stoel, R.B., 2005. Focus in Manado Malay. CNWS Publications, Leiden.

Thorsen, N., 1978. An acoustical analysis of Danish intonation. J. Phonet. 6, 151–175.

van Heuven, V.J., Haan, J., 2000. Phonetic correlates of statement versus question intonation in Dutch. In: Botinis, A. (Ed.), Intonation: Analysis, Modelling and Technology. Kluwer, Dordrecht, pp. 119–144.

van Heuven, V.J., Haan, J., 2002. Temporal development of interrogativity cues in Dutch. In: Gussenhoven, C., Warner, N. (Eds.), Laboratory Phonology 7. Mouton de Gruyter, Berlin, pp. 61–86.

van Heuven, V.J., van Leyden, K., 2003. A contrastive acoustical investigation of Orkney and Shetland intonation. In: Proc. 15th Internat. Congress of Phonetic Sciences, Barcelona, pp. 805–808.

van Leyden, K., 2004. Prosodic characteristics of Orkney and Shetland dialects. An experimental approach, LOT dissertation series 92, LOT, Utrecht.

van Zanten, E., Damen, L., van Houten, E., 1991. The ASSP database, SPIN/ASSP-report 41, Stichting Spraaktechnologie, Utrecht.

van Zanten, E., Damen, L., van Houten, E., 1993. Collecting data for a speech database. In: van Heuven, V.J., Pols, L.C.W. (Eds.), Analysis and synthesis of speech. Strategic research towards high-quality text-to-speech generation, Mouton de Gruyter, Berlin, pp. 207–222.

Vincent J. van Heuven (1949) obtained a BA and MA in English Language and Literature, an MA in General Linguistics, and a PhD in Psycholinguistics (all Utrecht University). He had a postgraduate year at Edinburgh University and a postdoc year at UCLA. He is professor of Experimental Linguistics and Phonetics at Universiteit Leiden, and acting director of the Universiteit Leiden Centre for Linguistics (ULCL). He is also vice-president of the Permanent Council of the International Phonetic Association (IPA).

Ellen van Zanten (1940) obtained a BA in Classical Languages (Utrecht University), an MA in General Linguistics and wrote a PhD dissertation on the vowels in varieties of Indonesian (higher degrees at Leyden University). She taught at Universitas Indonesia, and is now a senior researcher at the ULCL phonetics laboratory.