An unfamiliar intonation contour slows down on-line speech comprehension

Bettina Braun¹,², Audra Dainora³ and Mirjam Ernestus⁴,²

¹University of Konstanz, Germany
²Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands,
³Massachusetts Institute of Technology, Cambridge, MA
⁴Radboud University Nijmegen, The Netherlands

Running head: Unfamiliar intonation contours

Address for correspondence:
Bettina Braun
University of Konstanz
Department of Linguistics
Fach 186, D-78467 Konstanz, Germany
Email: Bettina.Braun@uni-konstanz.de
Tel: 0049-7531 882386
Fax: 0049-7531 882741
ABSTRACT

This study investigates whether listeners’ familiarity with an intonation contour affects speech processing. In three experiments, Dutch participants heard Dutch sentences with normal intonation contours and with unfamiliar ones and performed word monitoring, lexical decision, or semantic categorization tasks (the latter two with cross-modal identity priming). The unfamiliar intonation contour slowed down participants on all tasks, which demonstrates that an unfamiliar intonation contour has a robust detrimental effect on speech processing. Since cross-modal identity priming with a lexical decision task taps into lexical access, this effect obtained in this task suggests that an unfamiliar intonation contour hinders lexical access. Furthermore, results from the semantic categorization task show that the effect of an uncommon intonation contour is long-lasting and hinders subsequent processing. Hence, intonation not only contributes to (pragmatic) utterance meaning (emotion, sentence type, focus), but affects crucial aspects of the speech comprehension process and is more important than previously thought.

Keywords:

Intonation, speech comprehension, lexical access, cross-modal priming
INTRODUCTION

Every utterance, even if it consists only of a single word, is produced with a certain speech melody or intonation (see for instance the ‘Nine ways of saying yes’ in Crystal, 1995). Typologically, languages either use intonation for lexical purposes or not. In tone languages (e.g., Mandarin Chinese) and, to a lesser extent in pitch accent languages (e.g., Japanese, Swedish), pitch rises and falls distinguish between otherwise identical words (e.g., Hulst & Smith, 1988; Hyman, 1978). In intonation languages, intonation contributes to the (pragmatic) meaning of the whole sentence. More specifically, intonation signals attitudes and emotions (e.g., Liberman, 1975; Liberman & Sag, 1974; Scherer, Ladd, & Silverman, 1984), differentiates between sentence types such as statements and questions (e.g., Heuven & Haan, 2002), may indicate syntactic constituency (e.g., Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Snedeker & Trueswell, 2002; Speer, Kjelgaard, & Dobroth, 1996) conveys the domain of focus (e.g., Birch & Clifton, 2002; Welby, 2003) and marks contextually old and new information (e.g., Baumann, Grice, & Steindamm, 2006; Bock & Mazzella, 1983; Braun, 2006; Cutler & Foss, 1977; Dahan, Tanenhaus, & Chambers, 2002; Gussenhoven, 1984; Kohler, 1991; Terken & Nooteboom, 1987)

Notably, languages differ in their intonational realizations (e.g., Bolinger, 1978; Ladd, 1996), and so do dialects (see e.g. articles in Gilles & Peters, 2004). These differences can be phonological or phonetic in nature. For instance, Belfast English with its large number of high ending statements differs phonologically from most other dialects on the British Isles, in which statements end in low tones (Grabe, 2004). Northern and Southern German dialects, on the other hand, differ in the phonetic implementation of pitch accents: accentual peaks are aligned later with respect to the onset of the syllable in one dialect than in the other (e.g., Atterer & Ladd, 2004; Braun,
Furthermore, varieties of a language may differ in the relative frequencies of their contours (Dainora, 2006; Fletcher, Grabe, & Warren, 2005).

Acquiring a different intonation pattern in a second language (or dialect) is a difficult endeavor (e.g., Banjo, 1979; Cruz-Ferreira, 1989; Trouvain & Gut, 2007; Wennerstrom, 1994; Willems, 1982), so most speakers end up with an intonational foreign accent (e.g., Els & Bot, 1987; Jilka, 2000) even though they are otherwise highly proficient second language speakers. The flip side of the coin is that listeners are often exposed to intonation contours that do not exist in their own language variety and that they have heard rarely or never before.

Previous research has shown that it is difficult to comprehend speech with unfamiliar characteristics, such as the productions of speakers with a foreign accent (Bürki-Cohen, Miller, & Eimas, 2001) or of utterances with misplaced prosodic phrasing (e.g., Tyler & Warren, 1987). Only a few studies have investigated the effect of intonational foreign accent on speech comprehension. Their results suggest that also an unfamiliar intonation contour hinders comprehension. Holm (2007) studied the intelligibility of Norwegian as a second language by speakers of Dutch and English, among others. These non-native speakers’ productions were manipulated to match the intonational or durational properties of Norwegian as produced by native speakers. Results of an orthographic transcription task with native speakers of Norwegian showed that the corrected intonation enhanced intelligibility compared to the original productions. Munro and Derwing(1999) had English speakers rate foreign accent and comprehensibility of Mandarin speakers of English and measured the intelligibility with a transcription task. The authors coded the speech materials for phonemic, phonetic, and grammatical errors, as well as native-likeness of the intonation contour. Correlation analyses showed that of all these measures intonation correlated most strongly with
perceived foreign accent and comprehensibility. Intonation was also the second best predictor (after phonemic errors) for the numbers of errors in the transcription task. Finally, other studies have shown that intelligibility is reduced when the fundamental frequency is completely removed from a noise-masked sentence (Hillenbrand, 2003; Laures & Weismer, 1999).

All these studies suggest that an unfamiliar intonation contour may result in lower intelligibility, which indicates that intonation affects lexical access. However, listeners in the above described studies were presented with speech materials that had not only anomalous intonation but also anomalous segments (non-native or noise-masked). Furthermore, Holm (2007) and Munro and Derwing (1999) focused on speakers or learners of pitch accent or tone languages (Norwegian or Mandarin), in which pitch distinguishes between otherwise identical words, and thus constrains lexical access and selection (Cutler & Otake, 1999). Hence, it is still largely unknown whether in an intonation language an unfamiliar intonation of an otherwise completely native-like utterance affects lexical access.

Given that intonation in an intonation language contributes mostly to (pragmatic) utterance interpretation (e.g., sentence mode, attitudes, emotions), it appears more likely that intonation affects semantic integration (i.e. the combination of words into higher-level syntactic and semantic representations) rather than lexical access.

In this paper we investigated the role of intonation on speech comprehension in an intonation language. We tested whether there is an effect of an unfamiliar intonation contour on speech comprehension also in such a language and in the absence of other linguistic anomalies (non-native or noise-masked segments). Moreover we investigated specifically whether there is an effect of intonation on lexical access. Participants conducted three different tasks. Experiment 1 was a word-monitoring task (Kilburn &
Moss, 1996), which tested whether there is an effect of intonation when participants are encouraged to focus on the words of an utterance. Reaction times elicited with this experimental paradigm are determined both by the ease of lexical access and semantic integration. Hence, if there is any effect of an unfamiliar intonation contour on speech comprehension, it should surface with this task. Experiment 2 investigated directly whether an unfamiliar intonation contour affects lexical access. Participants performed a cross-modal identity priming experiment with a lexical decision task, which is argued to tap into a word’s lexical access (Norris, Cutler, McQueen, & Butterfield, 2006; Swinney, 1979). Finally, Experiment 3 was a cross-modal identity priming experiment with a semantic category judgment task, which follows lexical access. It allowed us to determine whether the effect of an uncommon intonation contour is robust, is long lasting, and also occurs when participants are focused on the lexical meaning of the words.

The experiments were conducted in Dutch with native speakers of Dutch. To ensure that all participants were equally unfamiliar with the contour used in our experiments, we generated one that is articulatorily possible, but – to our knowledge – unattested in any language. We generated a flipped sine-wave that slightly declined over time, with the declination to make it more speech-like (see Figure 1). The closest phonological transcription for this contour (according to ToDI, the Transcription of Dutch Intonation, as outlined in Gussenhoven, 2005) would be a falling prehead (%HL), followed by a half-completed fall (H*L%), which is associated with ‘Rome’ in the sentence shown in Figure 1. In the online course on Dutch intonation (http://todi.let.kun.nl/ToDI/home.htm), the falling prehead is explicitly described as being rare and its combination

---

1 There are a number of Dutch dialects that are described as tonal (i.e. in which pitch serves a lexical function, see e.g., Gussenhoven, 1999). Therefore, we only recruited participants who did not originate from areas with tonal dialects and who had no experience with a tone dialect or foreign tone language.
with a pitch accent is not discussed at all. Hereafter, we will refer to this contour as the “sine” contour.

An inexistent intonation contour is not comparable to an inexistent morpheme or word, for which no meaning can be retrieved from the mental lexicon. Recent experimental evidence suggests that unfamiliar intonation contours are mapped to the closest possible ’attractor’ or ’magnet’ (Braun, Kochanski, Grabe, & Rosner, 2006; Pierrehumbert & Steele, 1989), similar to what has been proposed for the perception of unfamiliar phonemes (Best, 1995; Best, Mc Roberts, & Goodell, 2001; Grieser & Kuhl, 1989; Kuhl, 1991). Pierrehumbert (Pierrehumbert & Steele, 1989) asked subjects to imitate a range of stimuli whose peak in fundamental frequency (the acoustic correlate of high pitch) was varied from an early to a late position in the syllable in several steps. Participants did not imitate the stimuli closely but produced accents with either an early or a late peak, which suggests that stimuli with intermediate alignments were mapped to the closest existing accent type. Braun et al., (2006) created sentences with hybrid intonation contours that lie between attested contours of English. Participants were asked first to imitate them and then to imitate their own imitations in three subsequent
sessions. Results showed that participants imitated some phonetic details, but over the course of the experiment, their productions collapsed towards attested English contours. Because of this mapping of unattested intonation contours to familiar ones, listeners may accommodate sentences with an inexistent contour (at least if these sentences sound natural, i.e., if the intonation is articulatorily possible and within a natural pitch range). As discussed in more detail below, this accommodation is also observed for the sentences with the sine contour: the speaker is interpreted mostly as unhappy or bored.

**EXPERIMENT 1**

Experiment 1 was a word monitoring experiment designed to test whether participants are slower in detecting a word when it is produced in a sentence with a sine intonation compared to a sentence with a normal intonation contour. Unlike the earlier intelligibility studies reported above, we manipulated only the intonation contour (via resynthesis), leaving segmental quality and speech rhythm unchanged. Experiment 1 therefore shows the unique contribution of listeners’ familiarity with the contour on speech comprehension.

**Materials**

We constructed 24 experimental sentences consisting of mostly sonorant sounds, which guaranteed a high resynthesis quality of the intonation contour (see below). For every sentence, we selected one target word to be monitored by the participant. The part-of-speech of the words was varied: In 14 of the experimental sentences, the target word was a content word (six nouns, four adjectives, and four verbs). In the remaining ten experimental sentences, participants had to monitor a function word (six prepositions
and four personal or possessive pronouns). As the word likelihood increases towards the end of sentences, the target words in our stimuli were mostly sentence-medial. The materials are listed in Appendix A.

In addition, we created 96 filler sentences with a different set of target words. The filler sentences consisted of both sonorant and non-sonorant sounds, but their syntactic and semantic structure was comparable to that of the experimental sentences. The part-of-speech categories of the filler target words were of the same relative frequency of occurrence as those of the experimental target words; there were 24 nouns, 16 adjectives, 16 verbs, 24 prepositions, and 16 pronouns. Like in the experimental trials, the target words were mostly sentence-medial. To keep participants attentive, thirty of the filler sentences did not contain the target word but a semantically related alternative (e.g., participants had to monitor the word “Amsterdam” in the sentence “The young girl goes for a walk in Suisse”). For these 30 sentences, participants had to monitor function or content words which were related mostly to words in sentence-medial position.

The sentences were recorded by a female speaker of Standard Dutch in a soundproof cabin. The speaker was not informed about the purpose of the experiment and not instructed which intonation contour to use. She read the sentences informally as they would be produced in spontaneous speech, which involved some high frequency reductions and elisions of sounds (e.g., elision of word-final /n/ after schwa in words such as *mennen* ’men’). On average overall duration of the sentences was 1.8 seconds (range: 1.3 to 2.2 seconds), which implies an average speech rate of 6 syllables per second. The experimental sentences were intonationally annotated by the first author and a student assistant trained in ToDI (Gussenhoven, 2005) Disagreements between the transcribers were discussed and resolved, and the resulting annotations are provided
in the Appendix. All six target nouns and two of the four target adjectives were accented while the remaining two adjectives, the six target verbs and all the function words were unaccented. We assume that these realizations are natural and represent the most common pronunciations of the sentences.

To create the unfamiliar sine intonation, sentences were resynthesised using the PSOLA-technique implemented in Praat (Boersma & Weenink, 1996). The resynthesis algorithm (see Appendix B) replaced the natural f0-contour of each experimental sentence with a time-warped declining inverse sine contour, generating f0-values in 5 ms steps. An example sine intonation is shown in Figure 1. The mean fundamental frequency and standard deviation of the sine intonation was matched to the original contour. In contrast to the natural contours, the f0-movements were not aligned with the syllable structure and the rises and falls were somewhat shallower in the sine contour. However, the sentences with the sine intonation had the same segmental, rhythmic, and prominence structure, and the same average fundamental frequency as the original sentences.

We also resynthesized the experimental sentences with a normal intonation contour by multiplying their f0-values by 1.05. Hence, all experimental sentences in Experiment 1 were created by resynthesis. For the sake of clarity, the multiplied intonation contour will nevertheless be referred to as the “normal” contour.

The materials were evaluated in two experiments conducted via the Web (for details see Appendix C). In the first Web experiment, participants rated the naturalness of sentences with a normal intonation contour, with the sine intonation contour and with a monotonous intonation contour. Results showed that the sine intonation contour was rated as significantly better than the monotonous one, albeit worse than the natural one. In the second Web experiment, participants heard sentences in a number of intonational
realizations (including the sine intonation) and had to indicate whether they heard a question or a statement and to describe the attitude of the speaker. Sentences with a sine intonation were mostly classified as sad or disappointed but agreement across participants was lower than for other intonation contours signaling extralinguistic meaning (surprised echo question and whining). Moreover, participants listened more often to the sentences with sine intonation than to the other sentences before they made their judgments.

**Participants**

Twenty-four native speakers of Dutch participated in the experiment. In this and in subsequent experiments, participants were paid a small fee, had no experience with tone dialects or languages, were naïve with respect to the purpose of the experiment, had no known hearing problems, and had not participated in the earlier experiments, evaluating the materials.

**Procedure**

Participants were tested individually in soundproof cabins sitting in front of a computer screen. Target words were presented in white Arial 72pt characters on a black background, and the sentences were played in stereo at a comfortable loudness via headphones.

Every trial started with the display of a star at the center of the screen. Then the target word appeared 300ms prior to the start of the spoken sentence. Participants received written instructions to press a button with their dominant hand as soon as they
heard the target word in the sentence. They were informed that some of the sentences
did not contain the target word and that they should not press the button in these cases.

Intonation was manipulated as a within-subjects, and for a given participant
between-items factor with 12 items per intonation condition per subject. The 24
experimental sentences were divided into two groups, with an equal number of nominal,
verbal, adjectival, prepositional, and pronominal target words. From these we created
two lists, each containing both groups of sentences but the sentences of one group were
presented with a normal intonation and the sentences of the other group with sine
intonation. The two lists together hence contained every sentence with both types of
intonation. In addition every list contained the 96 filler sentences. Six randomized
versions of the two lists were created (two participants for each list). Any two
experimental sentences with a sine contour were separated by at least four other
sentences. Fillers that did not contain the target word were likewise separated by at least
two intervening sentences.

The experiment consisted of 120 trials. The first five trials were fillers (including one
in which the target word did not appear in the sentence) and they were the same for all
participants. This first block was followed by a pause in which participants could ask
clarifying questions. The main body of the experiment started with two filler trials in
fixed order for all participants.

Reaction times were recorded relative to the auditory onset of the target word.
Button presses were only registered up to two seconds after the end of the auditory
sentence.
Results and Discussion

We excluded 21 target trials because there was no button press (timeout) and 10 because of negative reaction times (button presses before the acoustic onset of the target word). The reaction times of the remaining 545 trials were log transformed and subjected to a multi-level regression model with participant and target word as crossed random variables, and with contrast coding for factors (Baayen, 2008; Baayen, Davidson, & Bates, 2008; Bates, 2005; Pinheiro & Bates, 2000). The crucial predictor was the intonation of the sentence (normal vs. sine). Further, we included the part-of-speech of the target word as a predictor since function words may be harder to detect than content words; their meaning is dependent on the neighboring words and they are typically shorter, unaccented, and more reduced (e.g., Shi, Gick, Kanwischer, & Wilson, 2005). Finally, we included information about the position of the trial in the experiment (trial number) to account for tiredness or familiarization with the task.

Here and in the analyses of Experiments 2 and 3, we first tested for all main effects and interactions with intonation. We then removed predictors with a p-value larger than 0.1 if this did not deteriorate the fit of the model (as estimated by a log-likelihood test), but we kept all main effects for predictors that appeared in statistically significant interactions. Finally, the data points for which the absolute standardized residuals were greater than 2.5 were removed and the model was refitted. If predictors were not significant at the $p < 0.05$ level, they were removed and a new model was fitted. The final model, which we report here, consists only of significant predictors. P-values were

---

2 Lexical frequency might also affect reaction times but is strongly correlated with part-of-speech: the mean frequency for function words was 12.4 compared to 7.0 for content words ($t(22) = 9.4, p < 0.0001$, see Experiment 2 for how we determined these lexical frequencies). We therefore did not incorporate lexical frequency as a predictor.
estimated as the posterior probabilities of a Markov Chain Monte Carlo simulation with 10000 runs.

Table 1 reports the main effects and interactions of Experiment 1. The ‘:’ here and henceforth signals an interaction between the variables. Estimates and p-values are derived from a Markov Chain Monte Carlo (MCMC) simulation with 10000 runs. Positive estimates indicate the amount of increase in log-RT relative to the Intercept. For factors (e.g., part-of-speech), the change from the Intercept applies for the level given in italics (e.g., function word).

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimate (MCMC)</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>P (MCMC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (trial with a normal intonation and a content word target)</td>
<td>6.028</td>
<td>5.912</td>
<td>6.145</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>trial number</td>
<td>0.0009</td>
<td>0.0002</td>
<td>0.0016</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>part-of-speech (function word)</td>
<td>0.112</td>
<td>-0.041</td>
<td>0.257</td>
<td>n.s.</td>
</tr>
<tr>
<td>intonation (sine intonation)</td>
<td>0.057</td>
<td>0.001</td>
<td>0.109</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>intonation : part-of-speech</td>
<td>0.164</td>
<td>0.077</td>
<td>0.252</td>
<td>&lt;0.0005</td>
</tr>
</tbody>
</table>

Effect sizes are based on the statistical model given the median trial number (58).

In conclusion, performance is clearly slowed down by an unfamiliar intonation contour in word monitoring. Therefore, an uncommon intonation contour also hinders speech comprehension in an intonation language if all other characteristics of the
sentence (e.g., the phonemes) are undisturbed and natural. Our results extend earlier findings that sentences with foreign-accented or removed intonation contours are harder to understand (Hillenbrand, 2003; Holm, 2007; Laures & Weismer, 1999; Munro & Derwing, 1999).

Further, we found that the processing of function words is disturbed more by an unfamiliar intonation contour than the processing of content words. Function words are acoustically more reduced and suffer more from coarticulation than content words (e.g., Bell, Brenier, Gregory, Girand, & Jurafsky, 2009) and are hence more difficult to monitor. Since the processing of more difficult words is more sensitive to all kinds of factors known to influence word recognition (e.g., Feldman, Brown, & Pastizzo, 2006), it is not surprising that the effect of an unfamiliar intonation contour was larger for function words than for content words.

**EXPERIMENT 2**

The materials in Experiment 1 were generated via resynthesis, and participants may show special processing for such manipulated speech materials. In Experiment 2, we therefore tested the effect of an unfamiliar intonation contour with stimuli that were produced naturally. Furthermore, we changed the experimental paradigm to cross-modal identity priming with a lexical decision task, which has been shown to tap into lexical access. This experiment will therefore provide us with information about the locus of the effect.

Finally, the materials of Experiment 2 differed in their homogeneity from Experiment 1. Whereas Experiment 1 investigated the processing of function and target words in sentence medial and sentence final position, in Experiment 2, all experimental target words were content words, in sentence-final positions.
Materials

The same speaker as in Experiment 1 recorded twenty experimental sentences (listed in Appendix D), similar to the ones used in Experiment 1. As in Experiment 1, we automatically replaced the intonation contour of every target sentence by a sine intonation (with the same mean f0 and standard deviation as in the original recording). Then, our speaker listened to the resynthesised contours, looked at the f0 track, and imitated each of the resynthesised sentences three times. We calculated the RMS error between the sine resynthesis and every sine imitation for each sentence at 5 ms steps, and selected the imitation with the least RMS error. The mean RMS error across sentences was 30.5 (sd 12.5). RMS errors for each sentence are also provided in Appendix D. Figure 2 presents a sample f0-track of a sine imitation, together with the corresponding resynthesised sine contour.

The imitated sine intonation of Experiment 2 is very close to the resynthesised sine intonation used in Experiment 1. However, in contrast to Experiment 1, the rhythmic and prominence structure possibly differs between sentences with a normal intonation and sine intonation, but this is difficult to quantify and hence difficult to check. The sentences with an imitated sine intonation were 68.9 ms longer than sentences with the normal intonation contour (t(18) = 5.0, p < 0.001). More importantly, however, the prime words in the sentences with an imitated sine intonation and in the sentences with a normal intonation did not differ in duration (t(18) = 0.5, p > .5): The mean duration of these prime words was 449.3 ms in the sentences with an imitated sine intonation compared to 453.2 ms in the sentences with a normal intonation contour. All of them carried a pitch accent (mostly !H*L).
Further, we constructed 190 filler trials. Eighty-five of these filler trials had the prime word in sentence-final position and 105 in sentence-medial position. For 96 filler trials (i.e., half from the set with sentence-medial and half from the set with sentence-final primes), the visually presented target word was non-existent in Dutch (or other related languages such as English or German), but obeyed its phonotactic constraints. To prevent participants from developing a strategy of pressing the yes-button whenever they heard a sentence with an unusual intonation pattern, 16 of these filler sentences with non-word targets were resynthesised with a sine intonation (eight with final and eight with medial primes). In the remaining 94 filler trials, the visual target was an existing Dutch word, semantically unrelated to all the words in the sentence. The initial phonemes of both the word and non-word filler targets overlapped with the prime word (e.g., *winkel* ‘shop’ — *wikkel* ‘wrapper’). We opted for onset-overlapping filler targets to make the presence of identical prime-target pairs in the experimental trials less exceptional.

---

4 The proportion of sentences with the unfamiliar contour was kept low to avoid participants' familiarization with this contour.
Finally, we recorded 10 additional sentences, all with a normal intonation contour, for a familiarization phase. Half of them were combined with existing target words and the other half with non-word targets, which were semantically unrelated to the prime words (or any of the words in the sentence).

**Participants**

Twenty-eight native speakers of Dutch participated in the experiment. They had not participated in Experiment 1.

**Procedure**

A trial started with a star that was shown on the screen for 340 ms, followed by an auditorily presented sentence. The visual target word was presented half-way through the auditory prime word in the sentence and remained on screen until the end of the trial (two seconds after the end of the auditory stimulus).

Participants received written instructions that they should press the right button when the string of letters they saw on the screen was a Dutch word or proper name and the left button when it was not. The button-box was reversed for left-handed participants.

Intonation was manipulated as a within-subjects, between-items factor with 10 items in each of the two intonation conditions per subject. We created two experimental lists in the same way as for Experiment 1. In addition, every list contained the 190 filler sentences and 10 familiarization trials. The experiment totaled in 220 trials and contained two equally spaced breaks. Eight randomized versions of the two lists were constructed (3 participants for each list). The randomization procedure was similar to the one in Experiment 1. Participants were assigned randomly to a list.
In this and the following priming experiment, reaction times were measured from the presentation of the visual target. Responses occurring more than two seconds after the start of the visual display were not recorded.

**Results and Discussion**

Participants responded too slowly (slower than two seconds) in four experimental trials and responded incorrectly in another 12 experimental trials. The data for the remaining 545 experimental trials were analyzed in the same way as for Experiment 1, using multi-level regression models. The crucial predictor was sentence intonation (familiar vs. unfamiliar). Here and in the following priming experiment, control predictors were log-lexical frequency of the target word (calculated as the arithmetic mean between the log-lexical frequency of the word form as reported in the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995), based on large amounts of written texts, and on the log-lexical frequency in the smaller Corpus of Spoken Dutch (Oostdijk, et al., 2002)), its number of characters, the reaction time to the preceding filler trial (which is a strong predictor for the participant’s speed in that part of the experiment (e.g., De Vaan, Schreuder, & Baayen, 2007; Taylor & Lupker, 2001), and the position of the trial in the experiment.

We followed the fitting procedure for the statistical analysis as described in Experiment 1. The results are summarized in Table 2.
Table 2: Results for Experiment 2.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimate (MCMC)</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>P (MCMC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (trial with a normal intonation)</td>
<td>5.7397</td>
<td>5.3465</td>
<td>6.1249</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>reaction time to the preceding filler trial</td>
<td>0.1263</td>
<td>0.0689</td>
<td>0.1812</td>
<td>&lt;0.0005</td>
</tr>
<tr>
<td>trial number</td>
<td>-0.0005</td>
<td>-0.0007</td>
<td>-0.0002</td>
<td>0.0001</td>
</tr>
<tr>
<td>lexical frequency</td>
<td>-0.0259</td>
<td>-0.0413</td>
<td>0.0018</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>intonation (sine intonation)</td>
<td>0.0429</td>
<td>0.0123</td>
<td>0.0735</td>
<td>&lt;0.005</td>
</tr>
</tbody>
</table>

As summarized in Table 2, there was a main effect of intonation: responses were on average 27.4 ms slower when the sentence was produced with an unfamiliar intonation compared to a sentence with a familiar intonation (604.8 ms vs. 632.2 ms). The effects of the control predictors were comparable to those of Experiment 2.

This experiment replicated the effect of an unfamiliar intonation contour shown in Experiment 1. Since Experiment 2 uses natural productions of the unfamiliar contour, we can conclude that the effect is a genuine effect of the unfamiliar contour and not caused by segmental perturbations due to resynthesis.

More importantly, an unfamiliar intonation contour appears to slow down reaction times also in a cross-modal identity priming experiment with lexical decision. This suggests that intonation has a direct effect on lexical access. This is surprising, given that Dutch is an intonation language in which intonation is believed to mainly contribute to the (pragmatic) interpretation of an utterance.

To further explore the effect of an unfamiliar intonation contour, we replaced the lexical decision task with a semantic categorization in Experiment 3. With this task, reaction times are expected to be longer. If we find no effect of intonation, participants apparently overcome delays in lexical access due to unfamiliar intonation contours.

---

3 Effect sizes are calculated on the basis of the regression model, given the mean lexical frequency (4.9), the median trial number (112), following a trial with the mean RT (6.7).
rather quickly. In contrast, if we find an effect, the delays cannot be easily overcome or are strengthened by delay in semantic integration. The effect of an unfamiliar intonation contour is then robust and long lasting.

EXPERIMENT 3

Experiment 3 was a cross-modal identity priming experiment in which participants had to perform a semantic category judgment task (cf., De Jong, 2002; Landauer, Ross, & Didner, 1979). During the auditory presentation of a sentence, a visual target was presented on a computer screen. Participants had to indicate whether its referent was tangible or not.

Materials

The experimental materials and auditory filler sentences were identical to those in Experiment 2. All experimental sentences had tangible prime words in sentence-final position. For the filler trials, non-existing visual target words were replaced by non-tangible ones, existing ones by tangible ones.

Participants

Twenty-eight native speakers of Dutch participated in the experiment. They had not participated in the previous experiments.

Procedure

The procedure was similar to Experiment 2, but was adjusted to the semantic category judgment task. Participants were instructed to press the right button when the referent of
the visually presented word was tangible and the left button when it was not (the button box was reversed for left-handed participants). To make sure participants understood the definition of the semantic categories, we provided participants both with written instructions containing examples and with a short familiarization phrase of ten representative trials. During this familiarization phase, participants received feedback consisting of the display of the words “correct” or “incorrect” on the screen.

Intonation was manipulated as a within-subjects, between-items factor with 10 items for each of the two intonation conditions per subject. The experimental lists were identical to those in Experiment 2.

**Results and Discussion**

Fifteen experimental trials were discarded, six because of incorrect responses and nine because participants reacted too slowly (RTs longer than 2 seconds). Participants responded more slowly to experimental targets in this experiment than in Experiment 2 with a lexical decision task (the mean RT for Experiment 3 was 759 ms compared to 644 ms in Experiment 2).

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimate (MCMC)</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>P (MCMC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (trial with a normal intonation)</td>
<td>5.6673</td>
<td>5.1125</td>
<td>6.1938</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>reaction time to the preceding filler trial</td>
<td>0.1778</td>
<td>0.1033</td>
<td>0.2543</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>trial number</td>
<td>-0.0005</td>
<td>-0.0008</td>
<td>-0.0002</td>
<td>0.0005</td>
</tr>
<tr>
<td>lexical frequency</td>
<td>-0.0402</td>
<td>-0.0602</td>
<td>-0.0199</td>
<td>&lt;0.0005</td>
</tr>
<tr>
<td><strong>intonation</strong> (sine intonation)</td>
<td>0.0454</td>
<td>0.0077</td>
<td>0.0843</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>
The statistical analysis was identical to that of Experiment 2. As summarized in Table 3, there was a main effect of intonation: responses were on average 34.3 ms slower when the sentence was produced with an unfamiliar intonation compared to a familiar intonation (788.5 ms vs. 754.2 ms)\(^6\). The effects of the control predictors were comparable to those of Experiment 2. Additional analyses revealed that the effect of intonation contour was similar in Experiment 2 and 3 (no interaction between experiment and intonation contour in the combined data set).\(^7\)

To conclude, participants reacted more slowly upon hearing a sentence spoken with an unfamiliar intonation contour compared to a familiar one. The effect was comparable to that of Experiment 2 with a lexical decision task. This shows that the detrimental effect of an uncommon intonation contour is not a short-lived effect but has a longerlasting negative effect on speech comprehension.

**GENERAL DISCUSSION**

In this article we examined whether listeners of an intonation language such as Dutch (in which pitch variation does not serve a lexical function) process sentences with an unfamiliar intonation contour more slowly than those with more familiar contours. In particular, we tested whether this effect may be caused by difficulties in lexical access.

We conducted three experiments. Experiment 1 was a word-monitoring experiment with resynthesised materials (i.e. the sentences’ intonation was manipulated but not their segments). This paradigm is known to be sensitive to difficulties in lexical access and semantic integration. Experiment 2 was a cross-modal identity priming paradigm with a lexical decision task, which is especially sensitive to the speed of lexical access.

---

\(^6\)Effect sizes are calculated on the basis of the regression model, given the mean lexical frequency (4.9), the median trial number (112), following a trial with the mean RT (6.8).

\(^7\)Dividing the items into two bins (with fast and slow overall reaction times) did not interact with intonation (p > 0.5).
All stimuli in this experiment were produced naturally. Finally, Experiment 3 was identical to Experiment 2, except that, instead of performing lexical decision, participants had to categorize the target words as tangible or not. In all three experiments, participants reacted more slowly upon hearing a sentence with an unfamiliar intonation contour compared to a sentence with a familiar contour. The effects of an unfamiliar intonation contour in Experiment 2 and 3, which were highly similar, were comparable in size.

Our results extend previous research in several respects. We have shown that an unfamiliar intonation contour also affects speech processing in an intonation language and if the segments are undisturbed and unmasked by noise. In other words, the effect of an unfamiliar intonation contour is not limited to languages in which pitch variation serves a lexical function, and is not driven by segmental alternations which often accompany unfamiliar contours in the experiments reported in the literature (e.g., Holm, 2007; Munro & Derwing, 1999).

In addition, our results suggest that an unfamiliar intonation contour slows speech processing by affecting lexical access. The most direct evidence is provided by Experiment 2, which is based on the cross-modal identity priming paradigm combined with lexical decision. This paradigm is assumed to tap directly into lexical access (e.g., Norris, et al., 2006). The effect in Experiment 1 may be completely driven by lexical access as well, since participants were asked to monitor for words. Finally, the results of Experiments 3 may also be completely driven by lexical access. If so, we have to conclude that the effect of an unfamiliar intonation contour on lexical access is long lasting and that listeners do not easily recover from.

Participants performed a semantic categorization task in Experiment 3 and the unfamiliar contour may therefore have affected semantic processing here as well. Since
we did not find a difference in the size of the effect between Experiments 2 and 3, this would imply that in Experiment 3 participants had at least partly overcome the difficulties induced by the unfamiliar intonation on lexical access by the time they had performed semantic categorization. Independently of whether the effect of the unfamiliar contour in Experiment 3 is just caused by difficulties in lexical access or also by difficulties in semantic processing, the experiment clearly shows that the effect on speech comprehension is long lasting.

Note that our findings cannot simply be explained by the fact that the unfamiliar contour does not provide cues for speech segmentation. Prior work has shown that manipulating the alignment of pitch targets with the segmental structure alters segmentation strategies (e.g., Ladd & Schepman, 2003; Welby, 2007). In English, for instance, the exact position of a low tone before an accentual rise is a small aid in discriminating between pairs such as “Norma Nelson” and “Norman Elson” (Ladd & Schepman, 2003). However, it is unlikely that slower responses due to unfamiliar intonation contours are (solely) due to listeners’ difficulties with segmentation. First, in Experiment 1, we found that the effect of the unfamiliar intonation contour was greater for function words than for content words, even though these function words were all unstressed and therefore do not contain intonational cues to segmentation. Second, in Experiments 2 and 3, the sine contour was naturally produced by a native speaker of Dutch. While the resulting contour overall is very similar to an artificially generated sine contour, it is unlikely that the speaker could suppress or alter her usual intonational alignment cues to word segmentation. In other words, the overall contour was different for the sentences with a sine intonation than for the sentences with a normal intonation, but the two types of sentences may have been similar in their prosodic cues for segmentation, since both were produced by a native speaker.
The question then arises how to account for this effect of intonation on lexical access in an intonational language. One possibility is that intonation affects lexical access directly. This possibility is difficult to accommodate in psycholinguistic models of word comprehension that assume abstract representations for words, consisting of sequences of abstract phonemes, without information on intonation (e.g., Shortlist as described by Norris, 1994; Norris & McQueen, 2008). In such models, an unfamiliar contour has to affect the formation of prelexical representations, but it is unclear to us how. In contrast, a direct effect of an unfamiliar intonation can be more easily explained in exemplar models, which assume that words are lexically represented with all their redundant fine phonetic detail, including intonation (e.g., POLYSP, developed by Hawkins & Smith, 2001). An unfamiliar contour then results in some type of mismatch between the perceived word and the lexical representation of the word, which may lead to slower word recognition.

On the other hand, the effect of intonation on lexical access may also be more indirect. The evaluation of our materials has shown that listeners took more time to ascribe a meaning to utterances with the unfamiliar sine intonation. Furthermore, we know that intonational information is processed and interpreted as soon as it becomes available (e.g., Dahan, et al., 2002; Weber, Braun, & Crocker, 2006 and ). Consequently, the increased reaction times for stimuli with an unfamiliar intonation contour may result from increased difficulty in interpreting the intonation contour, which took resources from lexical access or diverted listeners' attention. Our results do not favor one interpretation over the other and further research is clearly necessary.

In conclusion, our study has shown that non-native intonation contours, even in the absence of non-native segmental pronunciations, delay lexical access, also in an intonational language. Language learners should try to acquire native-like intonation
contours, not only because non-native contours obviously mark them as foreigners, but also because such contours make them difficult to understand.
Appendix A

Experimental sentences in Experiment 1 (words to monitor in boldface). ToDI accent types are added in brackets.

1. De jonge\textsubscript{(H*L)} mannen waren in \textbf{Rome} \textsubscript{(H*L)}.
   \textit{The young men were in Rome.}

2. Wim \textsubscript{(L*H)} en Jan \textsubscript{(L*H)} wonen al jaren\textsubscript{(H*L)} in een \textbf{molen} \textsubscript{(H*L)}.
   \textit{Wim and Jan have lived in a mill for years now.}

3. Warner\textsubscript{(H*L)} neemt een jongen\textsubscript{(H*L)} mee naar de \textbf{woning} \textsubscript{(H*L)}
   \textit{Warner takes a boy with him to the house.}

4. Wij rijden\textsubscript{(H*L)} in mijn nieuwe\textsubscript{(H*)} \textbf{mini} \textsubscript{(H*L)} naar Arnhem \textsubscript{(H*L)}.
   \textit{We are driving to Arnhem in my new mini.}

5. Lara \textsubscript{(L*H)} en Julia\textsubscript{(H*)} willen mijn nieuwe\textsubscript{(H*L)} \textbf{juwelen} \textsubscript{(H*L)}.
   \textit{Lara and Julia want my new jewelry.}

6. Wij lijmen\textsubscript{(H*L)} een oranje\textsubscript{(H*L)} anjer \textsubscript{(L*H)} aan een lila \textsubscript{(L*H)} \textbf{mouw} \textsubscript{(H*L)}.
   \textit{We are gluing an orange carnation on a lilac sleeve.}

7. Mama\textsubscript{(H*L)} wil een \textbf{warme} melk\textsubscript{(H*L)}
   \textit{Mama wants a warm melk.}

8. Wij lenen\textsubscript{(H*L)} jouw \textbf{nieuwe}\textsubscript{(H*)} \textbf{emmer} \textsubscript{(H*L)}.
   \textit{We borrow your new bucket.}

9. Wij lenen \textsubscript{(L*H)} \textbf{zijn} \textsubscript{(H*L)} \textbf{nieuwe} \textsubscript{(H*L)} roman \textsubscript{(L*H)} aan Anja \textsubscript{(H*L)}.
   \textit{We lend his new novel to Anja.}

10. Meneer \textsubscript{(L*H)} en Li\textsubscript{(H*L)} zijn enorm\textsubscript{(H*)} \textbf{nare} mannen \textsubscript{(H*L)}.
    \textit{Mister La and Li are extremely horrible men.}
11. Oma (H*L) en oom Jan (H*L) naaien linnen (H*) mouwen (H*L).
   Grandmother and Uncle Jan sew linen sleeves.

12. Miljoenen (H*) mieren (H*L) wonen in mijn wei (H*L).
   Millions of ants live in my field.

13. Lia (H*L) en Marjolein (L*H) aaien een iele (H*) merel (H*L).
   Lia and Marjolein are petting a thin blackbird.

14. Mijn buren (H*L) mengen wijn (H*) en melk (H*L).
   My neighbours are mixing wine and milk.

15. Anne (H*L) en Mirjam (L*H) wenen in jouw armen (H*L).
   Anne and Mirjam are crying in your arms.

16. Manja (H*L) en Irene (L*H) eren Allah (H*L) in een arena (H*L).
   Manja and Irene praise Allah in an arena.

17. Een jongen (H*L) wil in een ruine (L*H) neurïen (H*L).
   A boy wants to hum in the ruins.

18. Een lawine (H*L) ramde in januari (L*H) een arena.
   An avalanche hit an arena in January.

19. Een merrie (L*H) en een reu (L*H) rennen (H*) om een weiland (H*L).
   A mare and a dog are running around a field.

20. Lea (H*L) en Ria (L*H) innen (H*L) mijn loon (H*) al jaren (H*L).
   Lea and Ria have collected my salary for years.

21. In mei (H*L) rennen wij naar Laren (H*L).
   In May we will run to Laren.

22. Jullie nemen (H*L) me mee (H*) naar Wenen (H*L).
   You are taking me to Vienna.

23. Jullie rouwen (H*L) om mijn arme (H*) oma (H*L).
You are grieving for my poor grandmother.

24. Er loeren reu(n) naar mijn lammeren (H*L).

Dogs are eyeing my lambs.
Appendix B

Algorithm to generate the sine contour for the individual experimental sentences.

sub create_sine
{

#frame size
$DT= 0.01;

#set number of f0-values:
$n=(round(($time_end-$time_start)/$DT))+1;

#empirically determined start-value
$startf0=$meanf0+22;

$deg=360;
$pi= 3.14;
$range=($sdf0*2);

for ($a=0; $a<=$deg; $a=$a+($deg/$n))
{

$f0=($startf0-($a/8))+$range*(cos($pi*($a+90)/180));

}

}
Appendix C

(a) Naturalness ratings

In the first web experiment, 18 participants rated the naturalness of the speech melody from 1 (very unnatural) to 5 (very natural) in (a) sentences with a normal (multiplied) intonation contour, (b) sentences with a sine intonation, and (c) sentences with a PSOLA resynthesised monotonous contour (declining at a similar rate as the sine intonation) functioning as a baseline. Participants heard each sentence only once, with one of the three intonation contours, and entered their rating via the keyboard. Mean rating was 1.7 for the flat contour, 2.2 for the sine intonation, and 4.1 for the normal contour. These results support our claim that the sine intonation contour is not very unnatural as it was rated significantly better than the monotonous intonation contour (p < 0.001, Wilcoxon signed ranks test). In addition, the difference in rating between the sine and the normal intonation contour shows that native Dutch speakers are not familiar with the sine contour.

(b) Interpretation of sentences with a sine contour

In the second Web experiment, we investigated whether sentences with a sine intonation are interpreted in a uniform way across listeners. The same female speaker provided five additional realizations of the 24 target sentences, two in which intonation provided a linguistic meaning (a neutral echo-question and a correction with a contrastive accent on one constituent) and three in which it provided extralinguistic meanings (a surprised echo-question, an excited exclamation, and a whining statement). Note that these realizations were play-acted and not controlled for voice quality, speech
rate, or intensity differences, all of which contribute to the expression of emotion and attitude (cf., (Williams & Stevens, 1972) (Scherer, 1986) (Mozzionacci, 1995) (Carlson, Granstrøm, & Nord, 1992) (Murray & Arnott, 1993) (Gobl & Chasaide, 2003)). The three extralinguistic meanings were signaled especially by voice quality.

Methods. Six experimental lists were constructed, each containing the 24 sentences in one of the five newly recorded realizations or with a sine intonation. Together these six lists contained all realizations of all sentences. Eighteen participants took part and were paid a small fee. They were randomly assigned to the six randomized lists. The participants’ task was to indicate for every sentence whether they heard a question or a statement and to describe the emotional state of the speaker. They were allowed to listen to the sound files as often as they wished. This number was recorded as an indication of task difficulty.

Results. Participants classified the neutral and surprised echo questions correctly as questions (82% and 89% of the cases) and the other realizations mostly as statements (correction: 91%, excited: 87%, whining: 99%, and sine contour: 100%). We categorized the perceived intention of the speaker (see Table 4 for the most frequent responses) into two broad classes, linguistic (e.g., question, statement, neutral) and extralinguistic (e.g., surprised, enthusiastic, angry). Participants sometimes provided several descriptions for one and the same trial and in these cases we only used the first description for the classification (e.g., “neutral, a bit disappointed” was classified as linguistic, whereas “a bit disappointed, neutral” was classified as extralinguistic).
Table 4: Results of Web experiment 2: Most frequent emotional categories (n >=4) for six different realizations of the 24 target sentences.

<table>
<thead>
<tr>
<th>Intended realization</th>
<th>Response</th>
<th>English translation</th>
<th>Number of occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>surprised echo question</td>
<td>verbaasd</td>
<td>surprised</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>vragend</td>
<td>questioning</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>neutral</td>
<td>neutral</td>
<td>8</td>
</tr>
<tr>
<td>neutral echo question</td>
<td>neutral</td>
<td>neutral</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>verbaasd</td>
<td>surprised</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>vragend</td>
<td>questioning</td>
<td>10</td>
</tr>
<tr>
<td>whining statement</td>
<td>verdrietig</td>
<td>sad</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>zielig</td>
<td>pitiful</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>wanhopig</td>
<td>desperate</td>
<td>7</td>
</tr>
<tr>
<td>exclamation</td>
<td>verbaasd</td>
<td>surprised</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>enthousiast</td>
<td>enthusiastic</td>
<td>5</td>
</tr>
<tr>
<td>correction</td>
<td>verbaasd</td>
<td>surprised</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>neutraal</td>
<td>neutral</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>verbeterend</td>
<td>correcting</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>ongeduldig</td>
<td>impatient</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>geirritteerd</td>
<td>irritated</td>
<td>5</td>
</tr>
<tr>
<td>sine intonation</td>
<td>verdrietig</td>
<td>sad</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>teleurgesteld</td>
<td>disappointed</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>neutral</td>
<td>neutral</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>mededeling</td>
<td>informing</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>ontevreden</td>
<td>disappointed</td>
<td>4</td>
</tr>
</tbody>
</table>

Participants tended to assign extralinguistic interpretations to all contours, including the linguistic ones (whining: 100%, excited: 92%, surprised: 74%, neutral echo question: 54%, correction: 62%, and sine: 80% of the cases). Importantly, the sine intonation contour elicited 80% responses of extralinguistic meaning, which is higher than for the contours with truly linguistic meanings (54% and 62%), but is also lower than expected if the sine contour just signals extralinguistic meaning. Moreover, participants did not agree so much about which extralinguistic meaning the sine intonation contour indicated: Whereas for the three contours with extralinguistic meanings there are words that were mentioned in at least 30% of cases (verdrietig ‘sad’

---

8 The surprised intonation elicited 9% “question” classifications, which explains the low percentage of responses of extralinguistic meaning assigned to this contour.
for the whining contour and *verbaasd* ‘surprised’ for the surprise intonation as well as for the excited contour), the most frequent interpretation for the sine contour, *verdrietig*, occurred only in 16% of trials. Most of the alternatives, however, appear to come from the same semantic domain (e.g., *teleurgестeld* ‘disappointed’, *triest* ‘sad’) and interestingly they all indicate some weak emotional state (e.g., the words *wanhopig* and *ontreddered*, both meaning ‘desparate’, were mentioned in no less than 12% of trials for the whining contours but never for the sine contours, even though *verdrietig* ‘sad’ was the most popular response for both contours). In summary, participants agree less on the sine intonation than on the existing intonations, as expected. Nevertheless, there appears to be some agreement about its interpretation.

Participants did not differ in how often they listened to the sine intonation and the neutral echo question (p > .1, Wilcoxon signed ranks test). However, they listened to the sine intonations more often than to the surprised echo questions (p < 0.05), to the whining realizations (p < 0.01), and the difference approached significance for the comparison between sine intonation and exclamation (p = 0.07) and between sine intonation and correction (p = 0.06). This suggests that although listeners can make sense of sentences with sine intonation, they need more time to interpret these sentences than for existing contours.
Appendix D

Experimental sentences in Experiment 2. Prime word was the sentence-final word.

(Numbers in brackets indicate RMS error between the sine resynthesis and the sine imitation, averaged in 5 ms steps).

1. Mama wil een warme melk.  
   *Mama wants a warm melk.*  
   
2. Wij lenen jouw rode emmer.  
   *We borrow your red bucket.*  
   
3. Miljoenen mieren wonen in mijn wei.  
   *Millions of ants live in my field.*  
   
   *We are driving to Arnhem in my new mini.*  
   
5. Lia en Marjolein aaien een iele merel.  
   *Lia and Marjolein are petting a thin blackbird.*  
   
6. La en Li zijn enorm nare mannen.  
   *Mister La and Li are extremely horrible men.*  
   
7. Een merrie en een reu rennen om een weiland.  
   *A mare and a dog are running around a field.*  
   
8. Wij lenen zijn nieuwe roman aan Anja.  
   *We lend his new novel to Anja.*  
   
   *Manja and Irene praise Allah in an arena.*  
   
10. De roeier joelden naar alle mensen aan de waalkade.  

36
The rower shouted at everyone on the Waalkade [riverside promenade in Nijmegen].

11. De jonge mannen waren in Rome.
   The young men were in Rome.

    Wim and Jan have lived in a mill for years now.

13. Wij lijmen een oranj anjer aan een lila mouw.
    We are gluing an orange carnation on a lilac sleeve.

    Anne and Mirjam are crying in your arms.

15. In mei rennen wij naar Laren.
    In May we will run to Laren.

    You are grieving for my poor grandmother.

17. Er loeren reu en naar mijn lammeren.
    Dogs are eyeing my lambs.

18. Twee oorwurmen waaie van een marmeren muur af.
    Two earwigs are blown off a marble wall.

19. De engelen jammeren om de lorrman en zijn zonen.
    The angels are wailing about the ragman and his sons.

20. Het hele najaar zeuren de jongens al om noren.
    The boys have been skating for the entire fall.


Acoustical Society of America, 93, 1097 – 1108.
and Speech, 46, 53-81.