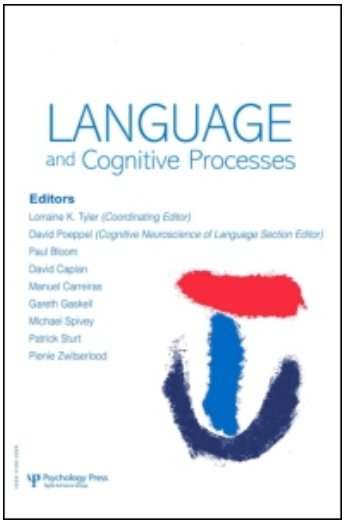
This article was downloaded by: *[Max Planck Institute for Psycholingistic]*
On: *15 March 2011*
Access details: *Access Details: [subscription number 930886218]*
Publisher *Psychology Press*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

An unfamiliar intonation contour slows down online speech comprehension

Bettina Braun[ab]; Audra Dainora[c]; Mirjam Ernestus[bd]
[a] Department of Linguistics, University of Konstanz, Konstanz, Germany [b] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands [c] Massachusetts Institute of Technology, Cambridge, MA, USA [d] Radboud University Nijmegen, Nijmegen, The Netherlands

First published on: 12 July 2010

To cite this Article Braun, Bettina , Dainora, Audra and Ernestus, Mirjam(2011) 'An unfamiliar intonation contour slows down online speech comprehension', Language and Cognitive Processes, 26: 3, 350 — 375, First published on: 12 July 2010 (iFirst)
To link to this Article: DOI: 10.1080/01690965.2010.492641
URL: http://dx.doi.org/10.1080/01690965.2010.492641

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

LANGUAGE AND COGNITIVE PROCESSES
2011, 26 (3), 350–375

Psychology Press
Taylor & Francis Group

An unfamiliar intonation contour slows down online speech comprehension

Bettina Braun[1,2], Audra Dainora[3], and Mirjam Ernestus[4,2]

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

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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 categorisation 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 categorisation task show that the effect of an uncommon intonation contour is long-lasting and hinders subsequent processing. Hence, intonation not only contributes to utterance meaning (emotion, sentence type, and focus), but also affects crucial aspects of the speech comprehension process and is more important than previously thought.

Keywords: Intonation; Speech comprehension; Lexical access; Cross-modal priming; Foreign accent; Dutch.

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

Correspondence should be addressed to Bettina Braun, Department of Linguistics, University of Konstanz, Fach 186, D-78467 Konstanz, Germany. E-mail: Bettina.Braun@uni-konstanz.de

© 2010 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business
http://www.psypress.com/lcp DOI: 10.1080/01690965.2010.492641

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

for lexical purposes or not. In tone languages (e.g., Mandarin Chinese) and to a lesser extent in pitch accent languages (e.g., Japanese and 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 realisations (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, 2007). 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 endeavour (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., van Els & de 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 number 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, and 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 (RTs) elicited with this experimental paradigm are determined both by the ease of lexical access and semantic integration. Hence, if there is any

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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 judgement task, which follows lexical access. It allowed us to determine whether the effect of an uncommon intonation contour is robust, 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.[1] 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. The declination was implemented 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 with a pitch accent is not discussed at all. Hereafter, we will refer to this contour as the "sine" contour.

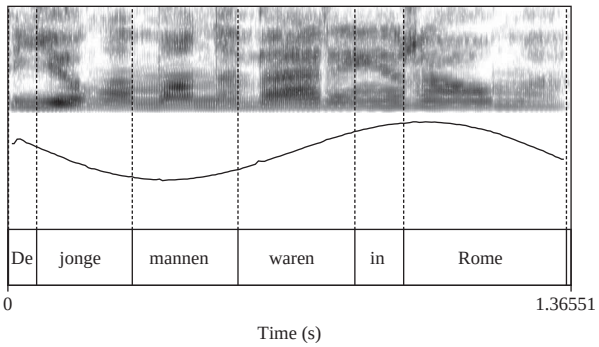


Figure 1. Unfamiliar sine contour for the sentence "The young men were in Rome".

[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.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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, McRoberts, & Goodell, 2001; Grieser & Kuhl, 1989; Kuhl, 1991). Pierrehumbert and 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

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

14 of the experimental sentences, the target word was a content word (six nouns, four adjectives, and four verbs). In the remaining 10 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 nonsonorant 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, 30 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 was 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 *mannen* "men"). On average overall duration of the sentences was 1.8 seconds (range: 1.3–2.2 seconds), which implies an average speech rate of six 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 Appendix A. 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 realisations 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 f_0 contour of each experimental sentence with a time-warped declining inverse sine contour, generating f_0 values in 5 ms steps. An example sine intonation is shown in Figure 1. The mean fundamental frequency and standard deviation (SD) of the sine intonation were matched to the original contour. In contrast to the natural contours, the f_0 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 resynthesised the experimental sentences with a normal intonation contour by multiplying their f_0 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 realisations (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 a sad or disappointed statement but agreement across participants was lower than for other intonation contours signalling 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 judgements.

Participants

Twenty-four native speakers of Dutch participated in the experiment. In this and in subsequent experiments, participants were paid a small fee. They had no experience with tone dialects or languages, and were naïve with respect to the purpose of the experiment. They 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 centre of the screen. Then the target word appeared 300 ms 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.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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, 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 randomised 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.

RTs were recorded relative to the auditory onset of the target word. Button presses were only registered up to 2 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 RTs (button presses before the acoustic onset of the target word). The RTs 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 neighbouring words and they are typically shorter, unaccented, and more reduced (e.g., Shi, Gick, Kanwischer, & Wilson, 2005).[2] Finally, we included information about the position of the

[2] Lexical frequency might also affect RTs but is strongly correlated with part-of-speech: the mean frequency for function words was 12.4 compared to 7.0 for content words [$t(22) = 9.4$, $p < .0001$, see Experiment 2 for how we determined these lexical frequencies.] We therefore did not incorporate lexical frequency as a predictor.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

trial in the experiment (trial number) to account for tiredness or familiarisation 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 .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 standardised residuals were greater than 2.5 were removed and the model was refitted. If predictors were not significant at the $p < .05$ 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 estimated as the posterior probabilities of a Markov Chain Monte Carlo simulation with 10,000 runs.

Table 1 reports the main effects and interactions of Experiment 1. The intercept refers to trials with a normal intonation contour and content word targets. As shown in Table 1, there was a main effect of intonation contour and an interaction between intonation contour and part-of-speech. Participants were slowed down by 3 ms on average when monitoring a content word (436.9 ms vs. 439.2 ms, $p < .05$) and by 141.9 ms when monitoring a function word (436.9 ms vs. 578.8 ms, $p < .0005$).[3] In addition, participants' reactions became slower during the course of the experiment.

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.

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

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March

TABLE 1
Results for Experiment 1.

Predictors	*Estimate (MCMC)*	*Lower bound*	*Upper bound*	*p (MCMC)*
Intercept (*trial with a normal intonation and a content word target*)	6.028	5.912	6.145	<.0001
Trial number	0.0009	0.0002	0.0016	<.05
Part-of-speech (*function word*)	0.112	−0.041	0.257	*ns*
Intonation (*sine intonation*)	0.057	0.001	0.109	<.05
Intonation * part-of-speech (*sine intonation, function word*)	0.164	0.077	0.252	<.0005

An asterisk between two factors signals an interaction. Estimates and *p*-values are derived from a Markov Chain Monte Carlo (MCMC) simuation with 10,000 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)

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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 20 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 f_0 and SD as in the original recording). Then, our speaker listened to the resynthesised contours, looked at the f_0 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 provided in Appendix D. Figure 2 presents a sample f_0 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

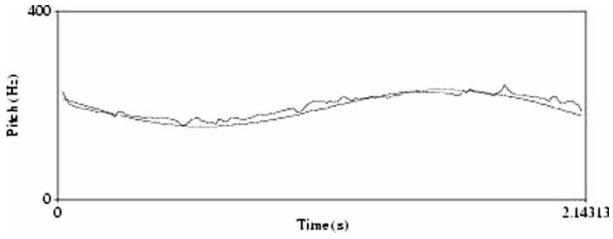


Figure 2. The f_0 track of sine imitation of the sentence "Lara en Julia willen mijn nieuwe juwelen" overlaid with the f_0 track of the artificial sine contour.

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, $t(18) = 5.0$, $p < .001$. 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 nonexistent 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 nonword targets were resynthesised with a sine intonation (eight with final and eight with medial primes).[4] 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 nonword 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.

Finally, we recorded 10 additional sentences, all with a normal intonation contour, for a familiarisation phase. Half of them were combined with existing target words and the other half with nonword 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 halfway through the auditory prime word in the sentence and remained on screen until the end of the trial (2 seconds after the end of the auditory stimulus).

[4] The proportion of sentences with the unfamiliar contour was kept low to avoid participants' familiarisation with this contour.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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 familiarisation trials. The experiment totalled in 220 trials and contained two equally spaced breaks. Seven randomised versions of the two lists were constructed (four participants for each list). The randomisation procedure was similar to the one in Experiment 1. Participants were assigned randomly to a list.

In this and the following priming experiment, RTs were measured from the presentation of the visual target. Responses occurring more than 2 seconds after the start of the visual display were not recorded.

Results and discussion

Participants responded too slowly (slower than 2 seconds) in four experimental trials and responded incorrectly in another 12 experimental trials. The data for the remaining 545 experimental trials were analysed 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 RT 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 summarised in Table 2.

As shown 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).[5] The effects of the control predictors were comparable to those of Experiment 2.

[5] 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).

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

TABLE 2
Results for Experiment 2

Predictors	*Estimate (MCMC)*	*Lower bound*	*Upper bound*	*p (MCMC)*
Intercept (*trial with a normal intonation*)	5.7397	5.3465	6.1249	<.0001
Reaction time to the preceding filler trial	0.1263	0.0689	0.1812	<.0005
Trial number	−0.0005	−0.0007	−0.0002	.0001
Lexical frequency	−0.0259	−0.0413	0.0018	<.005
Intonation (*sine intonation*)	0.0429	0.0123	0.0735	<.005

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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 RTs also in a cross-modal identity priming experiment with lexical decision, which suggests that intonation has a direct effect on lexical access. This is surprising, given that Dutch is an intonation language in which intonation mainly contributes 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 categorisation task in Experiment 3. With this task, RTs are expected to be longer. If we find no effect of intonation, participants apparently overcome delays in lexical access due to unfamiliar intonation contours 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 judgement 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, nonexisting visual target words were replaced by nontangible 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 judgement 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 familiarisation phrase of 10 representative trials. During this familiarisation 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).

The statistical analysis was identical to that of Experiment 2. As summarised 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 on the combined data set revealed that the effect of intonation contour was similar in Experiments 2 and 3 (no interaction between experiment and intonation contour).[7]

[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 RTs) did not interact with intonation ($p > 0.5$).

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

TABLE 3
Results for Experiment 3

Predictors	*Estimate (MCMC)*	*Lower bound*	*Upper bound*	*p (MCMC)*
Intercept (*trial with a normal intonation*)	5.6673	5.1125	6.1938	<.0001
Reaction time to the preceding filler trial	0.1778	0.1033	0.2543	<.0001
Trial number	−0.0005	−0.0008	−0.0002	.0005
Lexical frequency	−0.0402	−0.0602	−0.0199	<.0005
Intonation (*sine intonation*)	0.0454	0.0077	0.0843	<.05

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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 longer-lasting 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. 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 categorise 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 Experiments 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 when 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. Furthermore, the effect 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 Experiment 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 categorisation 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 categorisation. 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

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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 intonation 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). Consequently, the increased RTs 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 favour 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.

Manuscript received 5 September 2008
Revised manuscript received 7 May 2010
First published online 12 July 2010

REFERENCES

Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of "segmental anchoring" of *F0*: Evidence from German. *Journal of Phonetics*, *32*(3), 177–197.

Baayen, H. R. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

Baayen, H. R., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baayen, H. R., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database [CD-ROM]: Linguistic Data Consortium*. Philadelphia, PA: University of Pennsylvania.

Banjo, A. (1979). Beyond intelligibility in Nigerian English. In E. Ubahakwe (Ed.), *Varieties and functions of English in Nigeria* (pp. 7–13). Ibadan, Nigeria: African University Press.

Bates, D. (2005). Fitting linear mixed models in R. *R News*, *5*(1), 27–30.

Baumann, S., Grice, M., & Steindamm, S. (2006, May). *Prosodic marking of focus domains-categorical or gradient?* Paper presented at the the 3rd conference on speech prosody, Dresden, Germany.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*, 92–111.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore, MD: York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listeners' native phonological system. *Journal of the Acoustical Society of America*, *109*, 775–794.

Birch, S., & Clifton, C. (2002). Effects of varying focus and accenting of adjuncts on the comprehension of utterances. *Journal of Memory and Language*, *47*, 571–588.

Bock, K. J., & Mazzella, J. R. (1983). Intonational marking of given an new information: Some consequences for comprehension. *Memory & Cognition*, *11*(1), 64–76.

Boersma, P., & Weenink, D. (1996). *Praat, a system for doing phonetics by computer, version 3.4* (Report 132). Amsterdam: Institute of Phonetic Sciences, University of Amsterdam.

Bolinger, D. (1978). Intonation across languages. In J. Greenberg (Ed.), *Universals of Human Language, vol II: Phonology* (pp. 471–524). Palo Alto, CA: Stanford University Press.

Braun, B. (2006). Phonetics and phonology of thematic contrast in German. *Language and Speech*, *49*(4), 451–493.

Braun, B. (2007, August). *Effects of dialect and context on the realisation of prenuclear accents in German.* Paper presented at the XVIth international congress of the phonetic sciences, Saarbrücken, Germany.

Braun, B., Kochanski, G., Grabe, E., & Rosner, B. S. (2006). Evidence for attractors in English intonation. *Journal of the Acoustical Society of America*, *119*(6), 4006–4015.

Bürki-Cohen, J., Miller, J. L., & Eimas, P. D. (2001). Perceiving non-native speech. *Language and Speech*, *44*(2), 149–169.

Carlson, R., Granström, B., & Nord, L. (1992). Experiments with emotive speech, acted utterances and synthesized replicas. *Speech Communication*, *2*, 347–355.

Cruz-Ferreira, M. (1989). *Non-native comprehension of intonation patterns in Portuguese and in English*. Manchester, UK: University of Manchester.

Crystal, D. (1995). *The Cambridge encyclopedia of the English language*. Cambridge, UK: Cambridge University Press.

Cutler, A., & Foss, D. J. (1977). On the role of sentence stress in sentence processing. *Language and Speech*, *20*, 1–10.

Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America*, *105*(3), 1877–1888.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*(2), 292–314.

Dainora, A. (2006). Modeling intonation in English: A probabilistic approach to phonological competence. In L. Goldstein, D. Whalen, & C. Best (Eds.), *Laboratory phonology 8* (pp. 107–132). Berlin and New York: Mouton de Gruyter.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

De Jong, N. H. (2002). *Morphological families in the mental lexicon*. Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics.

De Vaan, L., Schreuder, R., & Baayen, R. H. (2007). Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon*, *2*(1), 1–23.

Feldman, L. B., Brown, D. B., & Pastizzo, M. J. (2006). Semantic influences on morphological facilitation: Concreteness and family size. *The Mental Lexicon Inaugural Issue*, *1*, 59–84.

Fletcher, J., Grabe, E., & Warren, P. (2005). Intonational variation in four dialects of English: The high rising tune. In S-A. Jun (Ed.), *Prosodic typology and transcription—A unifed approach* (pp. 390–409). Oxford, UK: Oxford University Press.

Gilles, P., & Peters, J. (Eds.). (2004). *Regional variation in intonation*. Tübingen, Germany: Niemeyer.

Gobl, C., & Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, *40*, 189–212.

Grabe, E. (2004). Intonational variation in English. In P. Gilles & J. Peters (Eds.), *Regional variation in intonation* (pp. 9–13). Tübingen: Niemeyer.

Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology*, *25*, 577–588.

Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents*. Dordrecht, The Netherlands: Foris Publications.

Gussenhoven, C. (1999, December). *Tone systems in Dutch Limburgian dialects*. Paper presented at the symposium on cross-linguistic studies of tonal phenomena: Tonogenesis, typology, and related topics, Tokyo: Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Languages.

Gussenhoven, C. (2005). Transcription of Dutch intonation. In S-A. Jun (Ed.), *Prosodic typology and transcription: A unified approach* (pp. 118–145). Oxford, UK: Oxford University Press.

Hawkins, S., & Smith, R. (2001). Polysp: A polysystemic, phonetically-rich approach to speech understanding. *Italian Journal of Linguistics—Rivista di Linguistica*, *13*, 99–188.

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

Hillenbrand, J. M. (2003). Some effects of intonation contour on sentence intelligibility. *Journal of the Acoustical Society of America Journal*, *114*, 2338–2338.

Holm, S. (2007). The relative contributions of intonation and duration to intelligibility in Norwegian as a second language. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the XVIth international congress of the phonetic sciences* (pp. 1653–1656). Saarbrücken, Germany: Pierrot.

Hulst, H. V. D., & Smith, N. (1988). *Autosegmental studies on pitch accent*. Dordrecht: Foris.

Hyman, L.M. (1978). Tone and/or accent. In D. Napoli (Ed.), *Elements of tone, stress, and intonation* (pp. 1–20). Washington, DC: Georgetown University Press.

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent*. Unpublished doctoral dissertation. Institut für Maschinelle Sprachverarbeitung.

Kilburn, K., & Moss, H. (1996). Word monitoring. *Language and Cognitive Processes*, *11*(6), 689–694.

Kohler, K. (1991). Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology and semantics. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)*, *25*, 115–185.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*, 93–107.

Ladd, D. R. (1996). *Intonational phonology*. New York, NY: Cambridge University Press.

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics*, *31*, 81–112.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

Landauer, T., Ross, B., & Didner, R. (1979). *Processing visually presented single words: A reaction time analysis* (Tech. Rep.). Murray Hill, NJ: Bell Laboratories.

Laures, J. S., & Weismer, G. (1999). The effects of a flattened fundamental frequency on intelligibility at the sentence level. *Journal of Speech, Language, and Hearing Research*, *42*, 1148–1156.

Liberman, M. (1975). *The intonational system of English*. Unpublished doctoral dissertation, MIT, Cambridge, MA.

Liberman, M., & Sag, I. (1974, April). *Prosodic form and discourse function.* Paper presented at the tenth regional meeting of the Chicago Linguistics Society, IL.

Mozzionacci, S. (1995, August). *Pitch variations and emotions in speech.* Paper presented at the XIIth international congress of the phonetic sciences, Stockholm.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*, 285–310.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature of human vocal emotion. *Journal of the Acoustical Society of America*, *93*, 1097–1108.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.

Norris, D., Cutler, A., McQueen, J., & Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology*, *53*, 146–193.

Norris, D., & McQueen, J. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357–395.

Oostdijk, N., Goedertier, W., van Eynde, F., Boves, L., Martens, J., & Moortgat, M., et al. (2002, May). *Experiences from the spoken Dutch corpus project.* Paper presented at the third international conference on language resources and evaluation, Las Palmas.

Pierrehumbert, J. B., & Steele, S. A. (1989). Categories of tonal alignment in English. *Phonetica*, *46*, 181–196.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, *90*(6), 2956–2970.

Scherer, K. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, *99*, 143–165.

Scherer, K. R., Ladd, D. R., & Silverman, K. E. A. (1984). Vocal cues to speaker affect: Testing two models. *Journal of the Acoustical Society of America*, *76*, 1346–1356.

Shi, R., Gick, B., Kanwischer, D., & Wilson, I. (2005). Frequency and category factors in the reduction and assimilation of function words: EPG and acoustic measures. *Journal of Psycholinguistic Research*, *34*, 341–364.

Snedeker, J., & Trueswell, J. (2002). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*, 103–130.

Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguity. *Journal of Psycholinguistic Research*, *25*, 247–268.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.

Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 117–138.

Terken, J. M. B., & Nooteboom, S. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes*, *2*, 145–163.

Trouvain, J., & Gut, U. (Eds.). (2007). *Non-native prosody: Phonetic description and teaching practice*. Berlin: Mouton de Gruyter.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

Tyler, L. K., & Warren, P. (1987). Local and global structure in spoken language comprehension. *Journal of Memory and Language*, *26*, 638–657.
van Els, T., & de Bot, K. (1987). The role of intonation in foreign accent. *Modern Language Journal*, *71*(2), 147–155.
Weber, A., Braun, B., & Crocker, M. (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech*, *49*, 367–392.
Welby, P. (2003). Effects of pitch accent type and status on focus projection. *Language and Speech*, *46*, 53–81.
Welby, P. (2007). The role of early fundamental frequency rises and elbows in French word segmentationation. *Speech Communication*, *49*, 28–48.
Wennerstrom, A. (1994). Intonational meaning in English discourse: A study of non-native speakers. *Applied Linguistics*, *15*(4), 399–420.
Willems, N. J. (1982). *English intonation from a Dutch point of view*. Dordrecht, The Netherlands: Foris Publications.
Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America*, *52*, 1238–1250.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

APPENDIX A

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

1. De jonge$_{(H*L)}$ mannen waren in **Rome** $_{(!H*L)}$.
   The young men were in Rome.
2. Wim $_{(L*H)}$ en Jan $_{(L*H)}$ wonen al jaren$_{(H*L)}$ in een **molen** $_{(!H*L)}$.
   Wim and Jan have lived in a mill for years now.
3. Warner$_{(H*L)}$ neemt een jongen$_{(H*L)}$ mee naar de **woning** $_{(!H*L)}$.
   Warner takes a boy with him to the house.
4. Wij rijden$_{(H*L)}$ in mijn nieuwe$_{(H*)}$ **mini** $_{(!H*L)}$ naar Arnhem $_{(!H*L)}$.
   We are driving to Arnhem in my new mini.
5. Lara $_{(L*H)}$ en Julia$_{(H*)}$ willen mijn nieuwe$_{(H*L)}$ **juwelen** $_{(!H*L)}$.
   Lara and Julia want my new jewelry.
6. Wij lijmen$_{(H*L)}$ een oranje$_{(H*L)}$ anjer $_{(L*H)}$ aan een lila $_{(L*H)}$ **mouw** $_{(!H*L)}$.
   We are gluing an orange carnation on a lilac sleeve.
7. Mama$_{(H*L)}$ wil een **warme** melk$_{(H*L)}$.
   Mama wants a warm melk.
8. Wij lenen$_{(H*L)}$ jouw **nieuwe**$_{(H*)}$ emmer $_{(!H*L)}$.
   We borrow your new bucket.
9. Wij lenen $_{(L*H)}$ zijn **nieuwe**$_{(H*L)}$ roman $_{(L*H)}$ aan Anja $_{(!H*L)}$.
   We lend his new novel to Anja.
10. Meneer La $_{(L*H)}$ en Li$_{(H*L)}$ zijn enorm$_{(H*)}$ **nare** mannen $_{(!H*L)}$.
    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 reuen$_{(H*)}$ 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
{
$DT = 0.01; #frame size
        #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.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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 monotonuous intonation contour ($p < .001$, Wilcoxon signed-rank tests). 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 realisations 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 realisations 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. Carlson, Granström, & Nord, 1992; Gobl & Chasaide, 2003; Mozzionacci, 1995; Murray & Arnott, 1993; Scherer, 1986; Williams & Stevens, 1972). The three extralinguistic meanings were signalled especially by voice quality.

Methods

Six experimental lists were constructed, each containing the 24 sentences in one of the five newly recorded realisations or with a sine intonation. Together these six lists contained all realisations of all sentences. Eighteen participants took part and were paid a small fee. They were randomly assigned to the six randomised 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 realisations mostly as statements (correction: 91%; excited: 87%; whining: 99%; and sine contour: 100%). We categorised the perceived intention of the speaker (see Table 4 for the most frequent responses) into two broad classes, linguistic (e.g., question, statement, and neutral) and extralinguistic (e.g., surprised, enthusiastic, and 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 extra-linguistic).

Participants tended to assign extralinguistic interpretations to all contours, including the linguistic ones (whining: 100%; excited: 92%; surprised: 74%;[8] 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

[8] The surprised intonation elicited 9% "question" classifications, which explains the low percentage of responses of extralinguistic meaning assigned to this contour.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

TABLE 4
Results of Web Experiment 2: most frequent emotional categories ($n \geq 4$) for six different realisations of the 24 target sentences

Intended realisation	*Response*	*English translation*	*Number of occurrences*
Surprised echo question	verbaasd	Surprised	28
	vragend	Questioning	10
	neutral	Neutral	8
Neutral echo question	neutral	Neutral	22
	verbaasd	Surprised	11
	vragend	Questioning	10
Whining statement	verdrietig	Sad	40
	zielig	Pitiful	15
	wanhopig	Ddesperate	7
Exclamation	verbaasd	Surprised	30
	enthousiast	Enthusiastic	5
Correction	verbaasd	Surprised	9
	neutraal	Neutral	8
	verbeterend	Correcting	6
	ongeduldig	Impatient	5
	geirriteerd	Irritated	5
Sine intonation	verdrietig	Sad	14
	teleurgesteld	Disappointed	10
	neutral	Neutral	9
	mededeling	Informing	4
	ontevreden	Disappointed	4

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" 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., teleurgesteld "disappointed", triest "sad") and interestingly they all indicate some weak emotional state (e.g., the words wanhopig and ontreddered, both meaning "desperate", 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-rank tests). However, they listened to the sine intonations more often than to the surprised echo questions ($p < .05$), to the whining realisations ($p < .01$), and the difference approached significance for the comparison between sine intonation and exclamation ($p = .07$) and between sine intonation and correction ($p = .06$). 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.

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011

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$_{(H*L)}$.
   Mama wants a warm melk. (60.10)
2. Wij lenen jouw rode emmer $_{(!H*L)}$.
   We borrow your red bucket. (31.10)
3. Miljoenen mieren wonen in mijn wei $_{(!H*L)}$.
   Millions of ants live in my field. (28.84)
4. Wij rijden in mijn nieuwe mini naar Arnhem $_{(!H*L)}$.
   We are driving to Arnhem in my new mini. (45.62)
5. Lia en Marjolein aaien een iele merel $_{(!H*L)}$.
   Lia and Marjolein are petting a thin blackbird. (19.05)
6. La en Li zijn enorm nare mannen $_{(!H*L)}$.
   Mister La and Li are extremely horrible men. (19.41)
7. Een merrie en een reu rennen om een weiland $_{(!H*L)}$.
   A mare and a dog are running around a field. (32.54)
8. Wij lenen zijn nieuwe roman aan Anja $_{(!H*L)}$.
   We lend his new novel to Anja. (22.14)
9. Manja en Irene eren Allah in een arena$_{(H*L)}$.
   Manja and Irene praise Allah in an arena. (27.64)
10. De roeier joelden naar alle mensen aan de waalkade $_{(!H*L)}$.
    The rower shouted at everyone on the Waalkade [riverside promenade in Nijmegen]. (34.42)
11. De jonge mannen waren in Rome $_{(!H*L)}$.
    The young men were in Rome. (22.62)
12. Wim en Jan wonen al jaren in een molen $_{(!H*L)}$.
    Wim and Jan have lived in a mill for years now. (25.39)
13. Wij lijmen een oranje anjer aan een lila mouw $_{(!H*L)}$.
    We are gluing an orange carnation on a lilac sleeve. (26.24)
14. Anne en Mirjam wenen in jouw armen $_{(!H*L)}$.
    Anne and Mirjam are crying in your arms. (24.27)
15. In mei rennen wij naar Laren $_{(!H*L)}$.
    In May we will run to Laren. (29.95)
16. Jullie rouwen om mijn arme oma $_{(!H*L)}$.
    You are grieving for my poor grandmother. (30.14)
17. Er loeren reuen naar mijn lammeren $_{(!H*L)}$.
    Dogs are eyeing my lambs. (24.69)
18. Twee oorwurmen waaien van een marmeren muur $_{(!H*L)}$ af.
    Two earwigs are blown off a marble wall. (60.78)
19. De engelen jammeren om de lorreman en zijn zonen $_{(!H*L)}$.
    The angels are wailing about the ragman and his sons. (68.38)
20. Het hele najaar zeuren de jongens al om noren $_{(!H*L)}$.
    The boys have been skating for the entire fall. (20.70)

Downloaded By: [Max Planck Institute for Psycholingistic] At: 08:30 15 March 2011