

An unfamiliar intonation contour slows down on-line speech comprehension

Journal:	<i>Language and Cognitive Processes</i>
Manuscript ID:	PLCP-2008-8508.R1
Manuscript Type:	Original Paper
Date Submitted by the Author:	16-Dec-2009
Complete List of Authors:	Braun, Bettina; Max Planck Institute for Psycholinguistics Dainora, Audra; Massachusetts Institute of Technology Ernestus, Mirjam; Radboud University; Max Planck Institute for Psycholinguistics
Keywords:	Intonation, speech comprehension, foreign accent, cross-modal priming



1
2
3
4
5
6
7
8
9 An unfamiliar intonation contour slows down on-line speech comprehension
10
11
12
13
14
15

16 Bettina Braun^{1,2}, Audra Dainora³ and Mirjam Ernestus^{4,2}
17
18
19
20

21 ¹University of Konstanz, Germany
22

23 ²Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands,
24

25 ³Massachusetts Institute of Technology, Cambridge, MA
26

27 ⁴Radboud University Nijmegen, The Netherlands
28
29
30
31
32
33

34 **Running head:** Unfamiliar intonation contours
35
36
37
38

39 **Address for correspondence:**
40

41 Bettina Braun
42

43 University of Konstanz
44

45 Department of Linguistics
46

47 Fach 186, D-78467 Konstanz, Germany
48

49 Email: Bettina.Braun@uni-konstanz.de
50

51 Tel: 0049-7531 882386
52

53 Fax: 0049-7531 882741
54
55
56
57
58
59
60

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 & Nootboom, 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

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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 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 (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,

1
2
3
4
5
6
7 as well as native-likeness of the intonation contour. Correlation analyses showed that of
8
9 all these measures intonation correlated most strongly with perceived foreign accent and
10
11 comprehensibility. Intonation was also the second best predictor (after phonemic errors)
12
13 for the numbers of errors in the transcription task. Finally, other studies have shown that
14
15 intelligibility is reduced when the fundamental frequency is completely removed from a
16
17 noise-masked sentence (Hillenbrand, 2003; Laures & Weismer, 1999).
18
19

20
21 All these studies suggest that an unfamiliar intonation contour may result in lower
22
23 intelligibility, which indicates that intonation affects lexical access. However, listeners
24
25 in the above described studies were presented with speech materials that had not only
26
27 anomalous intonation but also anomalous segments (non-native or noise-masked).
28
29 Furthermore, Holm (2007) and Munro (1999) focused on speakers or learners of pitch
30
31 accent or tone languages (Norwegian or Mandarin), in which pitch distinguishes
32
33 between otherwise identical words, and thus constrains lexical access and selection
34
35 (Cutler & Otake, 1999). Hence, it is still largely unknown whether in an intonation
36
37 language an unfamiliar intonation of an otherwise completely native-like utterance
38
39 affects lexical access.
40
41
42
43

44
45 Given that intonation in an intonation language contributes mostly to (pragmatic)
46
47 utterance interpretation (e.g., sentence mode, attitudes, emotions), it appears more likely
48
49 that intonation affects semantic integration (i.e. the combination of words into higher-
50
51 level syntactic and semantic representations) rather than lexical access.
52
53

54
55 In this paper we investigated the role of intonation on speech comprehension in
56
57 an intonation language. We tested whether there is an effect of an unfamiliar intonation
58
59 contour on speech comprehension also in such a language *and* in the absence of other
60
linguistic anomalies (non-native or noise-masked segments). Moreover we investigated

1
2
3
4
5
6 specifically whether there is an effect of intonation on lexical access. Participants
7
8 conducted three different tasks. Experiment 1 was a word-monitoring task (Kilburn &
9
10 Moss, 1996), which tested whether there is an effect of intonation when participants are
11
12 encouraged to focus on the words of an utterance. Reaction times elicited with this
13
14 experimental paradigm are determined both by the ease of lexical access and semantic
15
16 integration. Hence, if there is any effect of an unfamiliar intonation contour on speech
17
18 comprehension, it should surface with this task. Experiment 2 investigated directly
19
20 whether an unfamiliar intonation contour affects lexical access. Participants performed a
21
22 cross-modal identity priming experiment with a lexical decision task, which is argued to
23
24 tap into a word's lexical access (Norris, Cutler, McQueen, & Butterfield, 2006;
25
26 Swinney, 1979). Finally, Experiment 3 was a cross-modal identity priming experiment
27
28 with a semantic category judgment task, which follows lexical access. It allowed us to
29
30 determine whether the effect of an uncommon intonation contour is robust, is long
31
32 lasting, and also occurs when participants are focused on the lexical *meaning* of the
33
34 words.
35
36
37
38
39
40

41
42 The experiments were conducted in Dutch with native speakers of Dutch¹. To ensure
43
44 that all participants were equally unfamiliar with the contour used in our experiments,
45
46 we generated one that is articulatorily possible, but – to our knowledge – unattested in
47
48 any language. We generated a flipped sine-wave that slightly declined over time, with
49
50 the declination to make it more speech-like (see Figure 1). The closest phonological
51
52 transcription for this contour (according to ToDI, the Transcription of Dutch Intonation,
53
54 as outlined in Gussenhoven (2005) would be a falling prehead (%HL), followed by a
55
56
57

58
59 ¹ There are a number of Dutch dialects that are described as tonal (i.e. in which pitch serves a lexical
60 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.

1
2
3
4
5
6 half-completed fall (H*L%), which is associated with 'Rome' in the sentence shown in
7
8
9 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
10
11 with a pitch accent is not discussed at all. Hereafter, we will refer to this contour as the
12
13 "sine" contour.
14
15
16
17

18
19
20
21
22 Please insert Figure 1 around here.
23
24
25
26

27
28 An inexistent intonation contour is not comparable to an inexistent morpheme or
29
30 word, for which no meaning can be retrieved from the mental lexicon. Recent
31
32 experimental evidence suggests that unfamiliar intonation contours are mapped to the
33
34 closest possible 'attractor' or 'magnet' (Braun, Kochanski, Grabe, & Rosner, 2006;
35
36 Pierrehumbert & Steele, 1989), similar to what has been proposed for the perception of
37
38 unfamiliar phonemes (Best, 1995; Best, Mc Roberts, & Goodell, 2001; Grieser & Kuhl,
39
40 1989; Kuhl, 1991). Pierrehumbert (Pierrehumbert & Steele, 1989) asked subjects to
41
42 imitate a range of stimuli whose peak in fundamental frequency (the acoustic correlate
43
44 of high pitch) was varied from an early to a late position in the syllable in several steps.
45
46
47 Participants did not imitate the stimuli closely but produced accents with either an early
48
49 or a late peak, which suggests that stimuli with intermediate alignments were mapped to
50
51 the closest existing accent type. Braun (2006) created sentences with hybrid intonation
52
53 contours that lie between attested contours of English. Participants were asked first to
54
55 imitate them and then to imitate their own imitations in three subsequent sessions.
56
57
58
59
60 Results showed that participants imitated some phonetic details, but over the course of

1
2
3
4
5
6
7 the experiment, their productions collapsed towards attested English contours. Because
8
9 of this mapping of unattested intonation contours to familiar ones, listeners may
10
11 accommodate sentences with an inexistent contour (at least if these sentences sound
12
13 natural, i.e., if the intonation is articulatorily possible and within a natural pitch range).
14
15 As discussed in more detail below, this accommodation is also observed for the
16
17 sentences with the sine contour: the speaker is interpreted mostly as unhappy or bored.
18
19
20
21
22
23
24

25 **EXPERIMENT 1**

26
27 Experiment 1 was a word monitoring experiment designed to test whether participants
28
29 are slower in detecting a word when it is produced in a sentence with a sine intonation
30
31 compared to a sentence with a normal intonation contour. Unlike the earlier
32
33 intelligibility studies reported above, we manipulated only the intonation contour (via
34
35 resynthesis), leaving segmental quality and speech rhythm unchanged. Experiment 1
36
37 therefore shows the unique contribution of listeners' familiarity with the contour on
38
39 speech comprehension.
40
41
42
43
44
45
46
47
48
49

50 **Materials**

51 We constructed 24 experimental sentences consisting of mostly sonorant sounds, which
52
53 guaranteed a high resynthesis quality of the intonation contour (see below). For every
54
55 sentence, we selected one target word to be monitored by the participant. The part-of-
56
57 speech of the words was varied: In 14 of the experimental sentences, the target word
58
59 was a content word (six nouns, four adjectives, and four verbs). In the remaining ten
60

1
2
3
4
5
6
7 experimental sentences, participants had to monitor a function word (six prepositions
8
9 and four personal or possessive pronouns). As the word likelihood increases towards the
10
11 end of sentences, the target words in our stimuli were mostly sentence-medial. The
12
13 materials are listed in Appendix A.
14

15
16 In addition, we created 96 filler sentences with a different set of target words. The
17
18 filler sentences consisted of both sonorant and non-sonorant sounds, but their syntactic
19
20 and semantic structure was comparable to that of the experimental sentences. The part-
21
22 of-speech categories of the filler target words were of the same relative frequency of
23
24 occurrence as those of the experimental target words; there were 24 nouns, 16
25
26 adjectives, 16 verbs, 24 prepositions, and 16 pronouns. Like in the experimental trials,
27
28 the target words were mostly sentence-medial. To keep participants attentive, thirty of
29
30 the filler sentences did not contain the target word but a semantically related alternative
31
32 (e.g., participants had to monitor the word “Amsterdam” in the sentence “The young
33
34 girl goes for a walk in Suisse”). For these 30 sentences, participants had to monitor
35
36 function or content words which were related mostly to words in sentence-medial
37
38 position.
39
40
41
42
43

44
45 The sentences were recorded by a female speaker of Standard Dutch in a soundproof
46
47 cabin. The speaker was not informed about the purpose of the experiment and not
48
49 instructed which intonation contour to use. She read the sentences informally as they
50
51 would be produced in spontaneous speech, which involved some high frequency
52
53 reductions and elisions of sounds (e.g., elision of word-final /n/ after schwa in words
54
55 such as *mannen* ‘men’). On average overall duration of the sentences was 1.8 seconds
56
57 (range: 1.3 to 2.2 seconds), which implies an average speech rate of 6 syllables per
58
59 second. The experimental sentences were intonationally annotated by the first author
60

1
2
3
4
5
6 and a student assistant trained in ToDI (Gussenhoven, 2005) Disagreements between
7
8 the transcribers were discussed and resolved, and the resulting annotations are provided
9
10 in the Appendix. All six target nouns and two of the four target adjectives were
11
12 accented while the remaining two adjectives, the six target verbs and all the function
13
14 words were unaccented. We assume that these realizations are natural and represent the
15
16 most common pronunciations of the sentences.
17
18

19
20 To create the unfamiliar sine intonation, sentences were resynthesised using the
21
22 PSOLA-technique implemented in Praat (Boersma & Weenink, 1996). The resynthesis
23
24 algorithm (see Appendix B) replaced the natural f₀-contour of each experimental
25
26 sentence with a time-warped declining inverse sine contour, generating f₀-values in 5
27
28 ms steps. An example sine intonation is shown in Figure 1. The mean fundamental
29
30 frequency and standard deviation of the sine intonation was matched to the original
31
32 contour. In contrast to the natural contours, the f₀-movements were not aligned with the
33
34 syllable structure and the rises and falls were somewhat shallower in the sine contour.
35
36 However, the sentences with the sine intonation had the same segmental, rhythmic, and
37
38 prominence structure, and the same average fundamental frequency as the original
39
40 sentences.
41
42
43
44
45
46

47 We also resynthesized the experimental sentences with a normal intonation
48
49 contour by multiplying their f₀-values by 1.05. Hence, all experimental sentences in
50
51 Experiment 1 were created by resynthesis. For the sake of clarity, the multiplied
52
53 intonation contour will nevertheless be referred to as the “normal” contour.
54
55

56 The materials were evaluated in two experiments conducted via the Web (for
57
58 details see Appendix C). In the first Web experiment, participants rated the naturalness
59
60 of sentences with a normal intonation contour, with the sine intonation contour and with

1
2
3
4
5
6
7 a monotonous intonation contour. Results showed that the sine intonation contour was
8
9 rated as significantly better than the monotonous one, albeit worse than the natural one.
10
11 In the second Web experiment, participants heard sentences in a number of intonational
12
13 realizations (including the sine intonation) and had to indicate whether they heard a
14
15 question or a statement and to describe the attitude of the speaker. Sentences with a
16
17 sine intonation were mostly classified as sad or disappointed but agreement across
18
19 participants was lower than for other intonation contours signaling extralinguistic
20
21 meaning (surprised echo question and whining). Moreover, participants listened more
22
23 often to the sentences with sine intonation than to the other sentences before they made
24
25 their judgments.
26
27
28
29
30
31
32

33 **Participants**

34
35 Twenty-four native speakers of Dutch participated in the experiment. In this and in
36
37 subsequent experiments, participants were paid a small fee, had no experience with tone
38
39 dialects or languages, were naïve with respect to the purpose of the experiment, had no
40
41 known hearing problems, and had not participated in the earlier experiments, evaluating
42
43 the materials.
44
45
46
47
48
49
50

51 **Procedure**

52
53 Participants were tested individually in soundproof cabins sitting in front of a computer
54
55 screen. Target words were presented in white Arial 72pt characters on a black
56
57 background, and the sentences were played in stereo at a comfortable loudness via
58
59 headphones.
60

1
2
3
4
5
6
7 Every trial started with the display of a star at the center of the screen. Then the
8
9 target word appeared 300ms prior to the start of the spoken sentence. Participants
10
11 received written instructions to press a button with their dominant hand as soon as they
12
13 heard the target word in the sentence. They were informed that some of the sentences
14
15 did not contain the target word and that they should not press the button in these cases.
16
17

18
19 Intonation was manipulated as a within-subjects, and for a given participant
20
21 between-items factor with 12 items per intonation condition per subject. The 24
22
23 experimental sentences were divided into two groups, with an equal number of nominal,
24
25 verbal, adjectival, prepositional, and pronominal target words. From these we created
26
27 two lists, each containing both groups of sentences but the sentences of one group were
28
29 presented with a normal intonation and the sentences of the other group with sine
30
31 intonation. The two lists together hence contained every sentence with both types of
32
33 intonation. In addition every list contained the 96 filler sentences. Six randomized
34
35 versions of the two lists were created (two participants for each list). Any two
36
37 experimental sentences with a sine contour were separated by at least four other
38
39 sentences. Fillers that did not contain the target word were likewise separated by at least
40
41 two intervening sentences.
42
43
44
45

46
47 The experiment consisted of 120 trials. The first five trials were fillers (including one
48
49 in which the target word did not appear in the sentence) and they were the same for all
50
51 participants. This first block was followed by a pause in which participants could ask
52
53 clarifying questions. The main body of the experiment started with two filler trials in
54
55 fixed order for all participants.
56
57
58
59
60

1
2
3
4
5
6
7 Reaction times were recorded relative to the auditory onset of the target word.
8
9 Button presses were only registered up to two seconds after the end of the auditory
10
11 sentence.
12
13

14 15 16 17 **Results and Discussion**

18
19 We excluded 21 target trials because there was no button press (timeout) and 10
20
21 because of negative reaction times (button presses before the acoustic onset of the target
22
23 word). The reaction times of the remaining 545 trials were log transformed and
24
25 subjected to a multi-level regression model with participant and target word as crossed
26
27 random variables, and with contrast coding for factors (Baayen, 2008; Baayen,
28
29 Davidson, & Bates, 2008; Bates, 2005; Pinheiro & Bates, 2000). The crucial predictor
30
31 was the intonation of the sentence (normal vs. sine). Further, we included the part-of-
32
33 speech of the target word as a predictor since function words may be harder to detect
34
35 than content words; their meaning is dependent on the neighboring words and they are
36
37 typically shorter, unaccented, and more reduced (e.g., Shi, Gick, Kanwischer, &
38
39 Wilson, 2005)². Finally, we included information about the position of the trial in the
40
41 experiment (trial number) to account for tiredness or familiarization with the task.
42
43
44
45
46
47
48
49

50
51 Here and in the analyses of Experiments 2 and 3, we first tested for all main effects
52
53 and interactions with intonation. We then removed predictors with a p-value larger than
54
55 0.1 if this did not deteriorate the fit of the model (as estimated by a log-likelihood test),
56
57

58
59 ² Lexical frequency might also affect reaction times but is strongly correlated with part-of-speech: the
60 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.

1
2
3
4
5
6
7 but we kept all main effects for predictors that appeared in statistically significant
8 interactions. Finally, the data points for which the absolute standardized residuals were
9 greater than 2.5 were removed and the model was refitted. If predictors were not
10 significant at the $p < 0.05$ level, they were removed and a new model was fitted. The
11 final model, which we report here, consists only of significant predictors. P-values were
12 estimated as the posterior probabilities of a Markov Chain Monte Carlo simulation with
13 10000 runs.
14
15
16
17
18
19
20
21
22
23
24

25 Please insert Table 1 around here.
26
27
28
29

30 Table 1 reports the main effects and interactions of Experiment 1. The intercept
31 refers to trials with a normal intonation contour and content word targets. As shown in
32 Table 1, there was a main effect of intonation contour and an interaction between
33 intonation contour and part-of-speech. Participants were slowed down by 3 ms on
34 average when monitoring a content word (436.9 ms vs. 439.2 ms, $p < 0.05$) and by
35 141.9 ms when monitoring a function word (436.9 ms vs. 578.8 ms, $p < 0.0005$)³. In
36 addition, participants' reactions became slower during the course of the experiment.
37
38
39
40
41
42
43
44
45

46 In conclusion, performance is clearly slowed down by an unfamiliar intonation
47 contour in word monitoring. Therefore, an uncommon intonation contour also hinders
48 speech comprehension in an intonation language if all other characteristics of the
49 sentence (e.g., the phonemes) are undisturbed and natural. Our results extend earlier
50 findings that sentences with foreign-accented or removed intonation contours are harder
51
52
53
54
55
56
57
58
59
60

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

1
2
3
4
5
6
7 to understand (Hillenbrand, 2003; Holm, 2007; Laures & Weismer, 1999; Munro &
8
9 Derwing, 1999).

10
11 Further, we found that the processing of function words is disturbed more by an
12
13 unfamiliar intonation contour than the processing of content words. Function words are
14
15 acoustically more reduced and suffer more from coarticulation than content words (e.g.,
16
17 Bell, Brenier, Gregory, Girand, & Jurafsky, 2009) and are hence more difficult to
18
19 monitor. Since the processing of more difficult words is more sensitive to all kinds of
20
21 factors known to influence word recognition (e.g., Feldman, Brown, & Pastizzo, 2006),
22
23 it is not surprising that the effect of an unfamiliar intonation contour was larger for
24
25 function words than for content words.
26
27
28
29
30
31
32

33 **EXPERIMENT 2**

34
35 The materials in Experiment 1 were generated via resynthesis, and participants may
36
37 show special processing for such manipulated speech materials. In Experiment 2, we
38
39 therefore tested the effect of an unfamiliar intonation contour with stimuli that were
40
41 produced naturally. Furthermore, we changed the experimental paradigm to cross-modal
42
43 identity priming with a lexical decision task, which has been shown to tap into lexical
44
45 access. This experiment will therefore provide us with information about the locus of
46
47 the effect.
48
49
50

51
52 Finally, the materials of Experiment 2 differed in their homogeneity from
53
54 Experiment 1. Whereas Experiment 1 investigated the processing of function and target
55
56 words in sentence medial and sentence final position, in Experiment 2, all experimental
57
58 target words were content words, in sentence-final positions.
59
60

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 f_0 and standard deviation 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 also 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 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).

1
2
3
4
5
6
7
8
9
10
11 Please insert Figure 2 around here.
12
13
14
15

16 Further, we constructed 190 filler trials. Eighty-five of these filler trials had the
17 prime word in sentence-final position and 105 in sentence-medial position. For 96 filler
18 trials (i.e., half from the set with sentence-medial and half from the set with sentence-
19 final primes), the visually presented target word was non-existent in Dutch (or other
20 related languages such as English or German), but obeyed its phonotactic constraints.
21 To prevent participants from developing a strategy of pressing the yes-button whenever
22 they heard a sentence with an unusual intonation pattern, 16 of these filler sentences
23 with non-word targets were resynthesised with a sine intonation (eight with final and
24 eight with medial primes)⁴. In the remaining 94 filler trials, the visual target was an
25 existing Dutch word, semantically unrelated to all the words in the sentence. The initial
26 phonemes of both the word and non-word filler targets overlapped with the prime word
27 (e.g., *winkel* 'shop' — *wikkel* 'wrapper'). We opted for onset-overlapping filler targets
28 to make the presence of identical prime-target pairs in the experimental trials less
29 exceptional.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 Finally, we recorded 10 additional sentences, all with a normal intonation contour,
50 for a familiarization phase. Half of them were combined with existing target words and
51 the other half with non-word targets, which were semantically unrelated to the prime
52 words (or any of the words in the sentence).
53
54
55
56
57
58
59

60

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

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.

Please insert Table 2 around here.

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

1
2
3
4
5
6 compared to a sentence with a familiar intonation (604.8 ms vs. 632.2 ms)⁵. The effects
7
8 of the control predictors were comparable to those of Experiment 2.
9

10
11 This experiment replicated the effect of an unfamiliar intonation contour shown in
12
13 Experiment 1. Since Experiment 2 uses natural productions of the unfamiliar contour,
14
15 we can conclude that the effect is a genuine effect of the unfamiliar contour and not
16
17 caused by segmental perturbations due to resynthesis.
18
19

20
21 More importantly, an unfamiliar intonation contour appears to slow down reaction
22
23 times also in a cross-modal identity priming experiment with lexical decision. This
24
25 suggests that intonation has a direct effect on lexical access. This is surprising, given
26
27 that Dutch is an intonation language in which intonation is believed to mainly contribute
28
29 to the (pragmatic) interpretation of an utterance.
30
31

32
33 To further explore the effect of an unfamiliar intonation contour, we replaced the
34
35 lexical decision task with a semantic categorization in Experiment 3. With this task,
36
37 reaction times are expected to be longer. If we find no effect of intonation, participants
38
39 apparently overcome delays in lexical access due to unfamiliar intonation contours
40
41 rather quickly. In contrast, if we find an effect, the delays cannot be easily overcome or
42
43 are strengthened by delay in semantic integration. The effect of an unfamiliar intonation
44
45 contour is then robust and long lasting.
46
47
48
49
50

51 **EXPERIMENT 3**

52
53 Experiment 3 was a cross-modal identity priming experiment in which participants had
54
55 to perform a semantic category judgment task (cf., De Jong, 2002; Landauer, Ross, &
56
57
58

59
60

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).

1
2
3
4
5
6
7 Didner, 1979). During the auditory presentation of a sentence, a visual target was
8
9 presented on a computer screen. Participants had to indicate whether its referent was
10
11 tangible or not.
12

13 14 15 16 17 **Materials**

18
19 The experimental materials and auditory filler sentences were identical to those in
20
21 Experiment 2. All experimental sentences had tangible prime words in sentence-final
22
23 position. For the filler trials, non-existing visual target words were replaced by non-
24
25 tangible ones, existing ones by tangible ones.
26
27

28 29 30 31 **Participants**

32
33 Twenty-eight native speakers of Dutch participated in the experiment. They had not
34
35 participated in the previous experiments.
36
37

38 39 40 41 **Procedure**

42
43 The procedure was similar to Experiment 2, but was adjusted to the semantic category
44
45 judgment task. Participants were instructed to press the right button when the referent of
46
47 the visually presented word was tangible and the left button when it was not (the button
48
49 box was reversed for left-handed participants). To make sure participants understood
50
51 the definition of the semantic categories, we provided participants both with written
52
53 instructions containing examples and with a short familiarization phrase of ten
54
55 representative trials. During this familiarization phase, participants received feedback
56
57 consisting of the display of the words “correct” or “incorrect” on the screen.
58
59
60

1
2
3
4
5
6
7 Intonation was manipulated as a within-subjects, between-items factor with 10 items
8
9 for each of the two intonation conditions per subject. The experimental lists were
10
11 identical to those in Experiment 2.
12

13 14 15 16 **Results and Discussion**

17
18 Fifteen experimental trials were discarded, six because of incorrect responses and nine
19 because participants reacted too slowly (RTs longer than 2 seconds). Participants
20 responded more slowly to experimental targets in this experiment than in Experiment 2
21 with a lexical decision task (the mean RT for Experiment 3 was 759 ms compared to
22 644 ms in Experiment 2).
23
24
25
26
27
28
29
30
31
32
33

34 Please insert Table 3 around here.
35
36
37
38

39 The statistical analysis was identical to that of Experiment 2. As summarized in
40 Table 3, there was a main effect of intonation: responses were on average 34.3 ms
41 slower when the sentence was produced with an unfamiliar intonation compared to a
42 familiar intonation (788.5 ms vs. 754.2 ms)⁶. The effects of the control predictors were
43 comparable to those of Experiment 2. Additional analyses revealed that the effect of
44 intonation contour was similar in Experiment 2 and 3 (no interaction between
45 experiment and intonation contour in the combined data set).
46
47
48
49
50
51
52
53
54

55 To conclude, participants reacted more slowly upon hearing a sentence spoken with
56 an unfamiliar intonation contour compared to a familiar one. The effect was comparable
57
58
59

60

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).

1
2
3
4
5
6
7 to that of Experiment 2 with a lexical decision task. This shows that the detrimental
8
9 effect of an uncommon intonation contour is not a short-lived effect but has a longer-
10
11 lasting negative effect on speech comprehension.
12

13 14 15 16 **GENERAL DISCUSSION**

17
18 In this article we examined whether listeners of an intonation language such as Dutch
19
20 (in which pitch variation does not serve a lexical function) process sentences with an
21
22 unfamiliar intonation contour more slowly than those with more familiar contours. In
23
24 particular, we tested whether this effect may be caused by difficulties in lexical access.
25
26

27
28 We conducted three experiments. Experiment 1 was a word-monitoring experiment
29
30 with resynthesised materials (i.e. the sentences' intonation was manipulated but not
31
32 their segments). This paradigm is known to be sensitive to difficulties in lexical access
33
34 and semantic integration. Experiment 2 was a cross-modal identity priming paradigm
35
36 with a lexical decision task, which is especially sensitive to the speed of lexical access.
37
38 All stimuli in this experiment were produced naturally. Finally, Experiment 3 was
39
40 identical to Experiment 2, except that, instead of performing lexical decision,
41
42 participants had to categorize the target words as tangible or not. In all three
43
44 experiments, participants reacted more slowly upon hearing a sentence with an
45
46 unfamiliar intonation contour compared to a sentence with a familiar contour. The
47
48 effects of an unfamiliar intonation contour in Experiment 2 and 3, which were highly
49
50 similar, were comparable in size.
51
52
53
54

55
56 Our results extend previous research in several respects. We have shown that an
57
58 unfamiliar intonation contour also affects speech processing in an intonation language
59
60 and if the segments are undisturbed and unmasked by noise. In other words, the effect

1
2
3
4
5
6
7 of an unfamiliar intonation contour is not limited to languages in which pitch variation
8
9 serves a lexical function, and is not driven by alternations in the segments which often
10
11 accompany unfamiliar contours in the experiments reported in the literature (e.g., Holm,
12
13 2007; Munro & Derwing, 1999).

14
15
16 In addition, our results suggest that an unfamiliar intonation contour slows speech
17
18 processing by affecting lexical access. The most direct evidence is provided by
19
20 Experiment 2, which is based on the cross-modal identity priming paradigm combined
21
22 with lexical decision. This paradigm is assumed to tap directly into lexical access (e.g.,
23
24 Norris, et al., 2006)). The effect in Experiment 1 may be completely driven by lexical
25
26 access as well, since participants were asked to monitor for words. Finally, the results of
27
28 Experiments 3 may also be completely driven by lexical access. If so, we have to
29
30 conclude that the effect of an unfamiliar intonation contour on lexical access is long
31
32 lasting and that listeners do not easily recover from.

33
34
35
36
37 Participants performed a semantic categorization task in Experiment 3 and the
38
39 unfamiliar contour may therefore have affected semantic processing here as well. Since
40
41 we did not find a difference in the size of the effect between Experiments 2 and 3, this
42
43 would imply that in Experiment 3 participants had at least partly overcome the
44
45 difficulties induced by the unfamiliar intonation on lexical access by the time they had
46
47 performed semantic categorization. Independently of whether the effect of the
48
49 unfamiliar contour in Experiment 3 is just caused by difficulties in lexical access or also
50
51 by difficulties in semantic processing, the experiment clearly shows that the effect on
52
53 speech comprehension is long lasting.

54
55
56
57
58 Note that our findings cannot simply be explained by the fact that the unfamiliar
59
60 contour does not provide cues for speech segmentation. Prior work has shown that

1
2
3
4
5
6
7 manipulating the alignment of pitch targets with the segmental structure alters
8
9 segmentation strategies (e.g., Ladd & Schepman, 2003; Welby, 2007). In English, for
10
11 instance, the exact position of a low tone before an accentual rise is a small aid in
12
13 discriminating between pairs such as “Norma Nelson” and “Norman Elson” (Ladd &
14
15 Schepman, 2003). However, it is unlikely that slower responses due to unfamiliar
16
17 intonation contours are (solely) due to listeners' difficulties with segmentation. First, in
18
19 Experiment 1, we found that the effect of the unfamiliar intonation contour was greater
20
21 for function words than for content words, even though these function words were all
22
23 unstressed and therefore do not contain intonational cues to segmentation. Second, in
24
25 Experiments 2 and 3, the sine contour was naturally produced by a native speaker of
26
27 Dutch. While the resulting contour overall is very similar to the sine contour, it is
28
29 unlikely that the speaker could suppress or alter her usual intonational alignment cues to
30
31 word segmentation. In other words, the overall contour was different for the sentences
32
33 with a sine than for the sentences with a normal intonation, but the two types of
34
35 sentences may have been similar in their prosodic cues for segmentation, since both
36
37 were produced by a native speaker.

38
39
40
41
42
43
44 The question then arises how to account for this effect of intonation on lexical
45
46 access in an intonational language. One possibility is that intonation affects access
47
48 directly. This possibility is difficult to accommodate in psycholinguistic models of word
49
50 comprehension that assume abstract representations for words, consisting of sequences
51
52 of abstract phonemes, without information on intonation (e.g., Shortlist as described by
53
54 Norris, 1994; Norris & McQueen, 2008). In such models, an unfamiliar contour has to
55
56 affect the formation of prelexical representations, but it is unclear to us how. In contrast,
57
58 a direct effect of an unfamiliar intonation can be more easily explained in exemplar
59
60

1
2
3
4
5
6
7 models, which assume that words are lexically represented with all their redundant fine
8
9 phonetic detail, including intonation (e.g., POLYSP, developed by Hawkins & Smith,
10
11 2001). An unfamiliar contour then results in some type of mismatch between the
12
13 perceived word and the lexical representation of the word, which may lead to slower
14
15 word recognition.
16

17
18 On the other hand, the effect of intonation on lexical access may also be more
19
20 indirect. The evaluation of our materials has shown that listeners took more time to
21
22 ascribe a meaning to utterances with the unfamiliar sine intonation. Furthermore, we
23
24 know that intonational information is processed and interpreted as soon as it becomes
25
26 available (e.g., Dahan, et al., 2002; Weber, Braun, & Crocker, 2006 and).
27
28 Consequently, the increased reaction times for stimuli with an unfamiliar intonation
29
30 contour may result from increased difficulty in interpreting the intonation contour,
31
32 which took resources from lexical access or diverted listeners' attention. Our results do
33
34 not favor one interpretation over the other and further research is clearly necessary.
35
36
37
38

39
40 In conclusion, our study has shown that non-native intonation contours, even in
41
42 the absence of non-native segmental pronunciations, delay lexical access, also in an
43
44 intonational language. Language learners should try to acquire native-like intonation
45
46 contours, not only because non-native contours obviously mark them as foreigners, but
47
48 also because such contours make them difficult to understand.
49
50
51
52
53
54
55
56
57
58
59
60

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 milk.
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.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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) neurien _(!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.

1
2
3
4
5
6
7 23. Jullie rouwen_(H*L) om **mijn** arme_(H*) oma _(!H*L).

8
9 *You are grieving for my poor grandmother.*

10
11 24. Er loeren reuen_(H*) naar **mijn** lammeren _(!H*L).

12
13 *Dogs are eyeing my lambs.*
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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. 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).

1
2
3
4
5
6
7 Note that these realizations were play-acted and not controlled for voice quality, speech
8 rate, or intensity differences, all of which contribute to the expression of emotion and
9 attitude (cf., (Williams & Stevens, 1972) (Scherer, 1986) (Mozzionacci, 1995) (Carlson,
10 Granström, & Nord, 1992) (Murray & Arnott, 1993) (Gobl & Chasaide, 2003)). The
11 three extralinguistic meanings were signaled especially by voice quality.
12
13
14
15
16
17

18 **Methods.** Six experimental lists were constructed, each containing the 24 sentences
19 in one of the five newly recorded realizations or with a sine intonation. Together these
20 six lists contained all realizations of all sentences. Eighteen participants took part and
21 were paid a small fee. They were randomly assigned to the six randomized lists. The
22 participants' task was to indicate for every sentence whether they heard a question or a
23 statement and to describe the emotional state of the speaker. They were allowed to listen
24 to the sound files as often as they wished. This number was recorded as an indication of
25 task difficulty
26
27
28
29
30
31
32
33
34
35
36

37 **Results.** Participants classified the neutral and surprised echo questions correctly as
38 questions (82% and 89% of the cases) and the other realizations mostly as statements
39 (correction: 91%, excited: 87%, whining: 99%, and sine contour: 100%). We
40 categorized the perceived intention of the speaker (see Table 5 for the most frequent
41 responses) into two broad classes, linguistic (e.g., question, statement, neutral) and
42 extralinguistic (e.g., surprised, enthusiastic, angry). Participants sometimes provided
43 several descriptions for one and the same trial and in these cases we only used the first
44 description for the classification (e.g., "neutral, a bit disappointed" was classified as
45 linguistic, whereas "a bit disappointed, neutral" was classified as extralinguistic).
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Please insert Table 5 around here.

Participants tended to assign extralinguistic interpretations to all contours, including the linguistic ones (whining: 100%, excited: 92%, surprised: 74%⁷, 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' 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.

⁷ The surprised intonation elicited 9% "question" classifications, which explains the low percentage of responses of extralinguistic meaning assigned to this contour.

1
2
3
4
5
6
7 Participants did not differ in how often they listened to the sine intonation and the
8
9 neutral echo question ($p > .1$, Wilcoxon signed ranks test). However, they listened to the
10
11 sine intonations more often than to the surprised echo questions ($p < 0.05$), to the
12
13 whining realizations ($p < 0.01$), and the difference approached significance for the
14
15 comparison between sine intonation and exclamation ($p = 0.07$) and between sine
16
17 intonation and correction ($p = 0.06$). This suggests that although listeners can make
18
19 sense of sentences with sine intonation, they need more time to interpret these sentences
20
21 than for existing contours.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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)

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
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)

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.
- 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*. Ibadan: 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). *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: York Press.
- Best, C. T., Mc Roberts, 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 and new

- 1
2
3
4
5
6 information: Some consequences for comprehension. *Memory & Cognition*,
7 11(1), 64-76.
8
9
- 10 Boersma, P., & Weenink, D. (1996). Praat, a system for doing phonetics by computer,
11 version 3.4. Report of the Institute of Phonetic Sciences of the University of
12 Amsterdam, Report, 132
13
14
- 15 Bolinger, D. (1978). Intonation across languages. In J. Greenberg (Ed.), *Universals of*
16 *Human Language, vol II : Phonology* (Vol. 2, pp. 471 - 524). Palo Alto, CA:
17 Stanford University Press.
18
19
- 20 Braun, B. (2006). Phonetics and phonology of thematic contrast in German. *Language*
21 *and Speech*, 49(4), 451-493.
22
23
- 24 Braun, B. (2007). *Effects of dialect and context on the realisation of prenuclear accents*
25 *in German*. Paper presented at the XVth International Congress of the Phonetic
26 Sciences, Saarbrücken.
27
28
- 29 Braun, B., Kochanski, G., Grabe, E., & Rosner, B. S. (2006). Evidence for attractors in
30 English intonation. *Journal of the Acoustical Society of America*, 119(6), 4006-
31 4015.
32
33
- 34 Bürki-Cohen, J., Miller, J. L., & Eimas, P. D. (2001). Perceiving non-native speech.
35 *Language and Speech*, 44(2), 149-169.
36
37
- 38 Carlson, R., Granström, B., & Nord, L. (1992). Experiments with emotive speech,
39 acted utterances and synthesized replicas. *Speech Communication*, 2, 347 – 355.
40
41
- 42 Cruz-Ferreira, M. (1989). *Non-Native Comprehension of Intonation Patterns in*
43 *Portuguese and in English*. University of Manchester.
44
45
- 46 Crystal, D. (1995). *The Cambridge Encyclopedia of the English Language*. Cambridge,
47 UK: Cambridge University Press.
48
49
- 50 Cutler, A., & Foss, D. J. (1977). On the role of sentence stress in sentence processing.
51 *Language and Speech*, 20, 1-10.
52
53
- 54 Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese.
55 *Journal of the Acoustical Society of America*, 105(3), 1877 – 1888.
56
57
- 58 Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference
59 resolution in spoken-language comprehension. *Journal of Memory and*
60 *Language*, 47(2), 292-314.

- 1
2
3
4
5
6
7 Dainora, A. (2006). Modeling intonation in English: A probabilistic approach to
8 phonological competence. In L. Goldstein, D. Whalen & C. Best (Eds.),
9 *Laboratory Phonology 8* (pp. 107 – 132). Berlin, New York: Mouton de
10 Gruyter.
11
12
13 De Jong, N. H. (2002). *Morphological Families in the Mental Lexicon*. Max Planck
14 Institute for Psycholinguistics, Nijmegen, The Netherlands.
15
16 De Vaan, L., Schreuder, R., & Baayen, R. H. (2007). Regular morphologically complex
17 neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon*,
18 2(1), 1-23.
19
20
21
22
23 Els, T. v., & Bot, K. d. (1987). The role of intonation in foreign accent. *Modern*
24 *Language Journal*, 71(2), 147-155.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- 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 Unified Approach (pp. 390 – 409). Oxford: Oxford University
Press.
- Gilles, P., & Peters, J. (Eds.). (2004). *Regional Variation in Intonation*. Tübingen:
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, Holland; Cinnaminson, N.J., U.S.A.: Foris Publications.
- Gussenhoven, C. (1999). *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

- 1
2
3
4
5
6
7 Asia and Africa, Tokyo University of Foreign Languages.
- 8 Gussenhoven, C. (2005). Transcription of Dutch Intonation. In S.-A. Jun (Ed.),
9 *Prosodic typology and transcription: A unified approach* (pp. 118-145). Oxford:
10 Oxford University Press.
- 11
12
13
14 Hawkins, S., & Smith, R. (2001). Polysp: a polysystemic, phonetically-rich approach to
15 speech understanding. *Italian Journal of Linguistics – Rivista di Linguistica* 13,
16 99-188.
- 17
18
19 Heuven, V. J. v., & Haan, J. (2002). Temporal development of interrogativity cues in
20 Dutch. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory*
21 *Phonology* 7 (pp. 61-86). Berlin: Mouton de Gruyter.
- 22
23
24 Hillenbrand, J. M. (2003). Some effects of intonation contour on sentence intelligibility.
25 *Journal of the Acoustical Society of America Journal*, 114, 2338 –2338.
- 26
27
28 Holm, S. (2007). The relative contributions of intonation and duration to intelligibility
29 in Norwegian as a second language. In J. Trouvain & W. J. Barry (Eds.),
30 *Proceedings of the XVIth International Congress of the Phonetic Sciences* (pp.
31 1653 – 1656). Dudweiler, Saarbrücken, Germany Pierrot.
- 32
33
34 Hulst, H. v. d., & Smith, N. (1988). *Autosegmental Studies on Pitch Accent*. Dordrecht:
35 Foris.
- 36
37
38 Hyman, L. M. (1978). Tone and/or accent. In D. Napoli (Ed.), *Elements of Tone, Stress,*
39 *and Intonation* (pp. 1 – 20). Washington: Georgetown University Press.
- 40
41
42 Jilka, M. (2000). The Contribution of Intonation to the Perception of Foreign Accent.
43 Unpublished doctoral dissertation. Institut für Maschinelle Sprachverarbeitung.
- 44
45
46 Kilburn, K., & Moss, H. (1996). Word monitoring. *Language and Cognitive Processes*,
47 11(6), 689-694.
- 48
49
50 Kohler, K. (1991). Terminal intonation patterns in single-accent utterances of German:
51 phonetics, phonology and semantics. *Arbeitsberichte des Instituts für Phonetik*
52 *und digitale Sprachverarbeitung der Universität Kiel (AIPUK)*, 25, 115-185.
- 53
54
55 Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect"
56 for the prototypes of speech categories, monkeys do not. *Perception &*
57 *Psychophysics*, 50, 93 – 107.
- 58
59
60 Ladd, D. R. (1996). *Intonational Phonology*. Cambridge [England]; New York, NY,

- 1
2
3
4
5
6
7 USA: Cambridge University Press.
- 8 Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents
9 in English: Experimental evidence. *Journal of Phonetics*, 31(81-112).
- 10
11 Landauer, T., Ross, B., & Didner, R. (1979). *Processing visually presented single*
12 *words: A reaction time analysis (Tech. Rep.)*. Murray Hill, NJ: Bell
13 Laboratories.
- 14
15 Laures, J. S., & Weismer, G. (1999). The effects of a flattened fundamental frequency
16 on intelligibility at the sentence level. *Journal of Speech, Language, and*
17 *Hearing Research*, 42, 1148 – 1156.
- 18
19 Liberman, M. (1975). *The Intonational System of English*. Unpublished doctoral
20 dissertation, MIT.
- 21
22 Liberman, M., & Sag, I. (1974). *Prosodic form and discourse function*. Paper presented
23 at the Papers from the Tenth Regional Meeting of the Chicago Linguistics
24 Society, IL.
- 25
26 Mozzionacci, S. (1995). *Pitch variations and emotions in speech*. Paper presented at the
27 XIIth international Congress of the Phonetic Sciences, Stockholm.
- 28
29 Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and
30 intelligibility in the speech of second language learners. *Language Learning*, 49,
31 285 - 310.
- 32
33 Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic
34 speech: A review of the literature of human vocal emotion. *Journal of the*
35 *Acoustical Society of America*, 93, 1097 – 1108.
- 36
37 Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition.
38 *Cognition*, 52, 189-234.
- 39
40 Norris, D., Cutler, A., McQueen, J., & Butterfield, S. (2006). Phonological and
41 conceptual activation in speech comprehension. *Cognitive Psychology*, 53, 146 -
42 193.
- 43
44 Norris, D., & McQueen, J. (2008). Shortlist B: A Bayesian model of continuous speech
45 recognition. *Psychological Review*, 115(2), 357 – 395.
- 46
47 Oostdijk, N., Goedertier, W., Eynde, F. v., Boves, L., Martens, J., & Moortgat, M.
48 (2002). *Experiences from the spoken dutch corpus project*. Paper presented at
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4
5
6
7 the Third International Conference on Language Resources and Evaluation, Las
8 Palmas.
9
- 10 Pierrehumbert, J. B., & Steele, S. A. (1989). Categories of tonal alignment in English.
11 *Phonetica*, 46, 181 – 196.
12
- 13 Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New
14 York: Springer.
15
- 16 Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of
17 prosody in syntactic disambiguation. *The Journal of the Acoustical Society of*
18 *America*, 90(6), 2956-2970.
19
- 20 Scherer, K. (1986). Vocal affect expression: A review and a model for future research.
21 *Psychological Bulletin*, 99, 143 – 165.
22
- 23 Scherer, K. R., Ladd, D. R., & Silverman, K. E. A. (1984). Vocal cues to speaker affect:
24 Testing two models. *Journal of the Acoustical Society of America*, 76, 1346 –
25 1356.
26
- 27 Shi, R., Gick, B., Kanwischer, D., & Wilson, I. (2005). Frequency and category factors
28 in the reduction and assimilation of function words: EPG and acoustic measures.
29 *Journal of Psycholinguistic Research*, 34, 341 – 364.
30
- 31 Snedeker, J., & Trueswell, J. (2002). Using prosody to avoid ambiguity: Effects of
32 speaker awareness and referential context. *Journal of Memory and Language*,
33 48, 103 – 130.
34
- 35 Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic
36 structure on the resolution of temporary syntactic closure ambiguity. *Journal of*
37 *Psycholinguistic Research*, 25, 247 – 268.
38
- 39 Swinney, D. (1979). Lexical access during sentence comprehension: (re)consideration
40 of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645 –
41 659.
42
- 43 Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion
44 account. *Journal of Experimental Psychology: Learning, Memory, and*
45 *Cognition*, 27, 117 – 138.
46
- 47 Terken, J. M. B., & Nootboom, S. (1987). Opposite Effects of Accentuation and
48 Deaccentuation on Verification Latencies for Given and New Information.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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.

Tyler, L. K., & Warren, P. (1987). Local and Global Structure in Spoken Language Comprehension. *Journal of Memory and Language*, 26, 638-657.

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 segmentation. *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: 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.

Table 1: Results for 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).

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

Table 2: Results for Experiment 2.

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

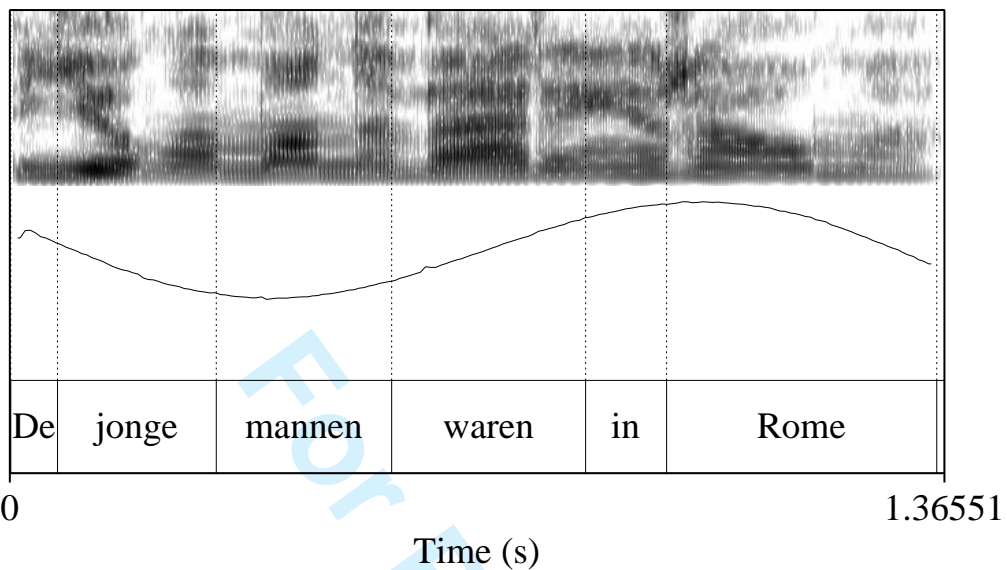
Table 3: Results for Experiment 3.

Predictors	Estimate (MCMC)	Lower bound	Upper bound	P (MCMC)
Intercept (<i>trial with a normal intonation</i>)	5.6673	5.1125	6.1938	<0.0001
reaction time to the preceding filler trial	0.1778	0.1033	0.2543	<0.0001
trial number	-0.0005	-0.0008	-0.0002	0.0005
lexical frequency	-0.0402	-0.0602	-0.0199	<0.0005
intonation (<i>sine intonation</i>)	0.0454	0.0077	0.0843	<0.05

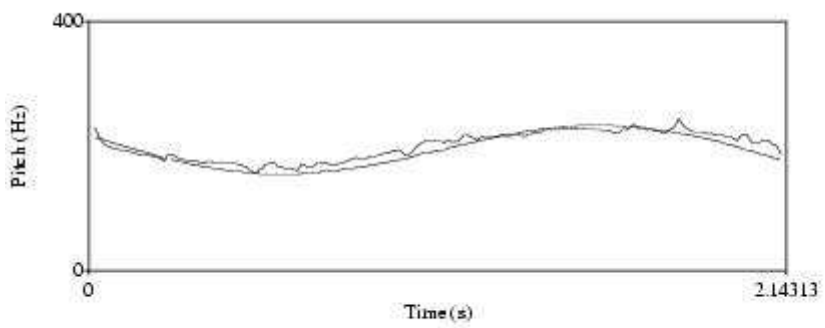
Table 5: Results of Web experiment 2: Most frequent emotional categories ($n \geq 4$) for six different realizations of the 24 target sentences.

Intended realization	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	desperate	7
exclamation	verbaasd	surprised	30
	enthousiast	enthusiastic	5
correction	verbaasd	surprised	9
	neutraal	neutral	8
	verbeterend	correcting	6
	ongeduldig	impatient	5
	geïrriteerd	irritated	5
sine intonation	verdrietig	sad	14
	teleurgesteld	disappointed	10
	neutral	neutral	9
	mededeling	informing	4
	ontevreden	disappointed	4

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



F0-track of sine imitation of an example sentence ("Lara en Julia willen mijn nieuwe juwelen") overlaid with the f0-track of the artificial sine contour.
165x76mm (72 x 72 DPI)

er Review Only