Laura Smorenburg 3843386
Hoefblad 20
4102 JP Culemborg
Supervisor: Aoju Chen
February 2015
Word Count: 7439

The Effect of Explicit Training on the Prosodic Production of L2 Sarcasm by Dutch Learners of English

BA Thesis

English Language and Culture

Utrecht University

Abstract

Previous research implied that Dutch learners of English are not able to accurately use prosody to convey sarcasm in English. The present study tested whether explicit instruction on the prosodic markers of sarcasm in English improved student's sarcastic sounding. English sarcastic speech was elicited from Dutch participants in two sessions. Between the two sessions, participants were trained in English sarcasm by (1) a presentation on the prosodic markers of sarcasm in English, (2) independent practice guided by audio and visual feedback on both target and own production and (3) an individual feedback session with the experimenters held in small groups. The degree of sarcasm in the recordings from the pre-test and post-test were evaluated by L1 British English-speaking raters on a five-point scale. Results showed a significant main effect of training, a significant main effect of sentence type as well as a significant interaction between the two. The findings imply that prosody training can have an effect on L2 learner's prosodic ability in the L2. In addition, some sentence types received higher average scores than others. Future research might explain the cause for the difference found between sentence types.

Table of Contents

Acknowledgements

It has become clear to me that one does not simply conduct an experiment on one's own, so I'd like to take this opportunity to express my thanks to my supervisor, Aoju Chen, and to Joe Rodd, who has played an invaluable part in this study. I really enjoyed being part of your project.

Secondly, I would like to thank all the people who participated in this study for their time and effort.

Lastly, I would like to acknowledge Diantha de Jong's work on her bachelor's thesis, which formed the point of departure for the present thesis.

1. Introduction

Much is still unknown about the prosodic patterns of languages. Prosody refers to the suprasegmental acoustic features of speech including rhythm, stress, and intonation. In recent years, more and more studies are conducted to shed light on prosody, but prosody from an interlinguistic perspective is hereby often overlooked. Additionally, even though correct use of prosody is important for effective communication (Weber et al., 2006), little attention is being paid to prosody in the teaching of second languages.

The present study is a follow-up on a study by De Jong (2014). She investigated whether Dutch learners of English could accurately make use of prosody to convey sarcasm in English. Pre-scripted sarcastic responses were elicited from participants via simulated conversations with a hypothetical 'friend'. British English-speaking raters as well as Dutch-speaking raters evaluated the sarcastic sounding of the responses on a five-point scale. De Jong's (2014) results imply that the Dutch learners of English in her study had not successfully adopted the British English prosodic system in their realisation of sarcasm, based on native-speaker ratings. In an L2 environment, the inability to accurately use prosody can have major consequences on communication (Brådvik et al., 1991). Correct use of prosody is especially important in sarcasm, because in sarcasm the speaker wishes to communicate the opposite of what is being said.

The current study has investigated whether it would be profitable for second language learners to receive explicit training on L2 prosodic systems by testing whether Dutch learners of English improve at conveying sarcasm via prosody after having received training on the English prosodic markers of sarcasm.

2. Previous Research on Sarcasm

2.1 Defining Sarcasm

To study prosodic production of sarcasm, firstly a definition of sarcasm has to be established, since the term can denote different meanings. Sarcasm is often confused with irony. The general definition of irony, as found in the English Oxford Dictionary, is expressing meaning by saying the opposite of what is intended, whereas sarcasm is more strictly defined as "the use of irony to mock or convey contempt". Anolli et al. (2002) suggest that sarcasm is a sub form of irony. They propose a main distinction of two forms of irony, namely "sarcastic irony" versus "kind irony". The first is defined as speech in which the literal reading is praising but the intended meaning is critical (1a) and the second is defined as speech in which the literal reading is critical but the intended meaning is complimentary (1b) (Anolli et al., 2002).

(1) *Literal reading*   *Intended meaning*

 a. I really like the colour of your dress. I really don't like the colour of your dress.

 b. You are the worst friend ever. You are the best friend ever.

Sarcasm can thus be characterised by its critical speaker intention as opposed to the praising intention in "kind irony". In perception, the distinction between "sarcastic irony" and "kind irony" remains arguable, as previous research has shown that the presence of a victim (as is the case in critical language) does not reliably predict the perception of sarcasm (Kreuz & Glucksberg, 1989). However, ratings for sarcasm were higher when there was a victim (p. 382), suggesting that, while the presence of a victim is not a requirement for the perception of sarcasm, it is a marker for sarcasm. This finding supports Anolli et al.'s (2002) suggestion

that sarcasm is a subtype of irony, and that "kind irony" as described by Anolli et al. (2002) is not sarcasm, because the speaker intent is praising (meaning the utterance does not contain a victim).

It has furthermore been suggested that speakers can control the weight of speaker intent in sarcasm, so that the covert critiques in sarcasm are either intensified (Jorgensen, 1996; Oring, 1994) or attenuated (Dews, Kaplan, & Winner, 1995). This suggests that the sarcastic sentence in (1a) can be uttered in a way that emphasises the critique as well as in a way that softens the critical speaker intent, resulting in a playful and a mocking type of sarcasm. Anolli et al. (2002) conducted an experiment to confirm the existence of the playful and mocking types of sarcasm as two separate forms of sarcasm. Participants were asked to respond with a given sentence that was placed in different contexts constructed to provoke either playful or mocking intent. Acoustic analysis revealed that two types of sarcasm could be discerned by their prosodic features.

In the present study sarcasm is considered to mean the playful version of "sarcastic irony". A single form was chosen because it has been shown by Anolli et al. (2002) that the different forms of sarcasm have different acoustic realisations. This variety is also acceptable in the broadest social context (Anolli et al., 2002).

2.2 Prosodic Markers of Sarcasm

Previous research has studied the perception of sarcasm in L1 by children versus adults and found that prosody plays a significant part in the perception of sarcasm from a young age (Laval, 2005; Capelli et al., 1990). Capelli et al. (1990) found that adults make use of both prosodic and contextual clues in perceiving sarcasm and that they can detect sarcasm equally well from either (p. 1836). Children aged eight, on the other hand, are better at perceiving sarcasm from prosody and do not seem to have the ability to rely on contextual clues

accurately, even in the presence of prosodic clues (Capelli et al., 1990, p. 1836). So prosody seems to be equally, if not more, important than context in the perception of sarcasm.

Previous findings on prosody of sarcasm include studies on the prosodic markers of sarcasm in Italian (Anolli et al., 2002) and English (Cheang & Pell, 2008, Attardo et al., 2003; Haiman, 1998). The varying results on the prosodic markers of sarcasm in English seem to suggest that Anolli et al's (2002) subtypes of sarcasm also apply in English. Anolli et al. (2002) found that the mocking type of sarcasm was realised with a "scornful voice" that is characterised by low mean pitch and little pitch variability in Italian, whereas high mean pitch, high pitch variability, increased loudness and reduced speech rate are considered characteristics of the playful "bantering voice" in Italian (p. 273-274).

The findings by Anolli et al. (2002) explain why different studies on sarcasm report different findings on the prosodic markers for sarcasm. For example, a study by Cheang & Pell (2008) on the prosodic markers of sarcasm in English found relatively low mean pitch and little pitch variability marked utterances for sarcasm (p. 372-374). Their definition of sarcasm emphasised that sarcasm was used to injure the listener by being mean and mocking (p. 366) and they thus seem to have stimulated the use of mocking sarcasm. Cheang & Pell's (2008) acoustic findings match Anolli et al's (2002) profile of the "scornful voice"[1] (p. 273). In contrast, other studies reported that sarcasm in English can be marked by large pitch variability, increased loudness, and slow articulation (Haiman, 1998, p. 33; Attardo et al. 2003), more closely matching the profile for the "bantering voice"[1] that Anolli et al. (2002) found for playful sarcasm.

No empirical data on the prosodic markers of sarcasm in Dutch can be found in existing literature. A small dataset from a bilingual speaker of British English and Dutch,

---

[1] Note that Anolli et al's (2002) findings were based on sarcastic production in Italian, not English.

which was collected in the present study, seems to suggest that sarcastic irony in Dutch might be marked by a wide pitch span and slower speech relative to neutral speech.

2.3 Other Markers of Sarcasm

As noted before in section 2.2, adult speakers (but not young speakers) can perceive sarcasm without the presence of acoustic markers for sarcasm by inferring sarcastic speaker intention from the context (Capelli et al., 1990). For example, when it's storming outside and someone remarks that *"the weather is beautiful today",* the listener might deduce from the context that the statement is counterfactual and conclude that the speaker is being sarcastic.

Other non-prosodic markers for sarcasm are kinesic markers, or facial expressions (Attardo et al., 2003). Examples are winks and certain eyebrow movements (Attardo et al., 2003, p. 245). These kinesic markers can be a way for the speaker to accentuate the sarcastic intent of an utterance, but exaggerated use of them is optional and they are therefore not always noticeably present in spontaneous speech (Attardo et al., 2003, p. 246).

Additionally, sarcastic speaker intent can be marked by the utterance itself through lexical cues. It has been suggested that some lexical units are so frequently used in a sarcastic manner that they have become key phrases for sarcasm (Haiman, 1998; Kreuz & Caucci, 2007). Clear examples of lexical items that have a sarcastic bias are English particles *"gee"* and *"gosh"* (Kreuz & Caucci, 2007, p. 2-3).

Certain syntactic structures may also be markers for sarcasm, for example, the tagquestion, which is a structure that is often used rhetorically (Kreuz & Caucci, 2007, p. 2). Because tag-questions are meant to be interpreted in a non-literal manner so frequently, these structures might be more susceptible to sarcastic interpretation than other structures.

The more markers for sarcasm are used, the easier it should be for listeners to perceive sarcasm. Earlier findings support this hypothesis, as utterances that had both prosodic and

contextual clues received relatively higher scores for sarcasm than when only one of the two was present (Capelli et al., 1990; Bryant & Tree, 2002).

2.4 Prosody of Sarcasm Across Languages

L2 prosody of sarcasm is a subject which has not yet received a lot of attention from researchers. Scholars argue that there is a universal "ironic tone of voice" (e.g., Creusere, 1999, p. 231), which presumes that irony is conveyed in the same manner emotions are, i.e., that irony is, or is associated with, an emotion. Emotions have certain physical effects on the body which in turn affect the voice (Cheang & Pell, 2011, p. 206; Haiman, 1998, p. 27). For example, an emotion causing a tensed body puts tension on the vocal tract, resulting in a strained voice. Cheang & Pell (2011) suspected that the acoustic markers for sarcasm were not universal and conducted a study to investigate whether sarcasm could be recognised in another language. They found that their results do not support the notion of a universal "ironic tone of voice", as participants could not accurately recognise sarcasm in the L2 for the language pairs English and Cantonese. They furthermore suggest that prosody of sarcasm cannot be perceived in a second language without additional non-prosodic cues (Cheang & Pell, 2011). These findings imply that sarcasm is not associated with one particular emotion and cannot be defined in terms of physical traits like emotions. Moreover, findings from previous research imply it is challenging for children to perceive sarcasm in a second language: children aged eight or older cannot use context accurately to detect sarcasm (Capelli et al., 1990) and, additionally, are suggested to be unable to perceive L2 prosody of sarcasm (Cheang & Pell, 2011). It is unclear whether adult L2 speakers can perceive sarcasm from contextual clues better than L1-speaking children can, but if not, it would seem that L2 sarcasm is highly difficult to perceive.

De Jong (2014) has tested whether Dutch learners of English were able to accurately use prosody to indicate sarcastic intent in their English speech, i.e.: whether they could sound sarcastic in the L2. Although there are some studies on the perception of L2 prosody of sarcasm, De Jong's (2014) study is the first to look at L2 prosodic production of sarcasm. She found that Dutch-speaking raters gave higher scores for sarcasm than English-speaking raters did. Her findings implicate that Dutch speakers could not accurately make use of prosody to convey sarcasm in English, as the speaker's sarcastic utterances seemed to contain (Dutch) prosodic markers for sarcasm that Dutch raters recognised, but English raters did not. De Jong (2014) furthermore argues that this variance in scores per rater-language supports Cheang & Pell's (2011) finding that prosodic markers of sarcasm differ across languages. In this case that would mean that the prosodic markers of sarcasm in Dutch are different from the markers in English.

In summary, not many studies have looked at the perception and production of interlingual sarcasm, but existing literature seems to suggest sarcasm is not easily perceived and produced in a second language.

2.5 Explicit Training in Prosodic Markers of Sarcasm

Explicit training has proven valuable in improving L2 learners' perception and production of certain prosodic patterns. However, it is unclear whether explicit training in the L2 prosodic system can improve second language learners' ability to produce sarcasm, since this has not yet been studied. As a consequence, formal education does not generally provide explicit instructions on prosody, let alone the prosody of sarcasm.

An example of a study that reported improved perception of prosodic patterns in L2 English after training is a study by Pincus (2014). The training consisted of a short instruction that linked negative versus positive speaker intent to a certain pitch contour for three

intonation patterns in English. Pincus (2014) implemented audio-visual techniques to aid participant's learning. Chinese participants listened to a recording of a native speaker and were asked to identify positive versus negative speaker intent by choosing one of the two follow-up sentences that were presented on-screen. An example of one of their stimuli is: *"Emily has beautiful hair"* (p. 3428). The object "hair" received a falling pitch or a rising-falling pitch accent to differentiate between neutral/positive and negative speaker intent respectively, which is not an acoustic parameter other studies have mentioned when describing the markers of sarcasm. The results showed that a short ten-minute training in between two experimental sessions improved participant's accuracy scores significantly (Pincus, 2014, p. 3433).

Another study tested whether Italian learners of German could improve their German pronunciation with training (Missaglia, 2007). More specifically, Missaglia (2007) looked at the effect of training in segmental features (now conventional in the teaching of second languages) versus training in prosodic features on Italian pronunciation by L2 learners. Segmental pronunciation training is the norm in the teaching of second languages. The results indicate that training in prosodic features of the L2 is more profitable than training in segmental features, as both the overall comprehensibility and segmental productions of L2 speakers that received only prosodic training improved relative to the control group that received only segmental-based training (Missaglia, 2007, p. 252). This study shows that prosodic training may have positive results on L2 learner's prosodic ability, based on L1 speakers' ratings.

3. Present Study

The present study was conducted to test whether prosody training can improve Dutch learners of English' accuracy in expressing sarcasm in English using prosody.

As mentioned in section 1, De Jong (2014) found that Dutch speakers of English are not very successful in adopting the English prosodic system in their realisation of sarcasm in the L2. Taking De Jong's findings (2014) as a point of departure, the current study investigates the effects of explicit prosodic training on the production of L2 sarcasm. The main research question is as follows:

RQ$_1$: Can explicit training in the English prosodic markers of sarcasm positively influence Dutch learners' accuracy in producing sarcasm in English?

We are furthermore interested in any differences for sentence types in the productions of sarcasm. De Jong (2014) found an effect for sentence type in her study; particle sentences were rated higher than all other sentence types, what-exclamations were rated slightly higher than declaratives, and there was no significant difference between declaratives and tag-questions. She hypothesised that tag-questions and particle sentences would receive higher scores, as those supposedly contain non-prosodic markers of sarcasm. This hypothesis was not supported by her results. Because the effect of sentence type is still unclear the possible differences in the accuracy in Dutch learners of English' ability to convey sarcasm between sentence types will also be investigated.

Additionally, training might have different effects in some sentence types than others. It might be the case that students improve more in some sentence types than in others. The last research question is as follows:

RQ$_2$: Do Dutch learners of English show more relative improvement in conveying sarcasm after prosody training in some sentence types than others?

Based on previous research on prosody training we expect a positive relation between the ratings of sarcastic production of Dutch learners of English and explicit training in the English prosodic markers of sarcasm, i.e.: an improvement in their sarcastic intonation following the prosody training ($\mu_{pre\text{-}test} < \mu_{post\text{-}test.}$). The following hypothesis is put forward:

H₁: Explicit training in the English prosodic markers of sarcasm will result in an increase in scores for sarcastic tone in the post-test relative to the pre-test.

Because De Jong's (2014) results on effect of sentence type were surprising and not easily explained by the theoretic framework, sentence type will again be investigated in the present study. It is predicted that structures that contain non-prosodic markers of sarcasm, such as tag-questions, will receive higher sarcasm sounding scores than structures that do not have non-prosodic markers of sarcasm, such as declaratives and what-exclamations.

Lastly, no differences are expected in the effect size of training between sentence types. This prediction is based on intuitions, as we are not aware of any research on interaction between different structures and effect size of training. The last hypothesis is:

H₂: Dutch learners of English will not show significantly larger improvement in some sentence types than others following prosody training.

4. Method

This study consists of three major parts; (1) a production experiment, (2) prosody training and (3) a rating experiment. In the first experiment, sarcastic production was elicited from Dutch

learners of English on two occasions: in week one (pre-test) and in week three (post-test). In between the test sessions participants received prosody training on the markers of sarcasm in English. In the rating experiment, the recordings from the pre- and post-test were rated by L1 British English-speakers on a five-point scale to evaluate the sarcastic sounding. A study by Carey et al. (2011) warns researchers that experiments that use L1 speakers to evaluate L2 speech are not always reliable. They found that evaluations of pronunciation are influenced by the rater's familiarity with the non-native accent. The British English rater's familiarity with Dutch L2 speech is, however, not a factor in the present study, because no statements about absolute proficiency are made. Rather, there is an interest in relative improvement of post-test scores, to look at the effect of training. Nevertheless, the group of raters was controlled for familiarity with the Dutch language, so that none of the raters were familiar with Dutch and presumably the Dutch accent in English.

4.1 Pre- and Post-Test

4.1.1 Participants

Twelve university students (9 females, 3 males, $M_{age} = 21.3$, SD = 1.7) participated in the pre- and post-test. They were all second or third-year students of the BA programme in English Language and Culture at Utrecht University and were estimated to have an advanced level of British English proficiency[2] (C1/C2 on the Common European Framework of References for Languages (Beeker et al., 2010)). All participants attended the course Prosody of English and Dutch at Utrecht University. At the moment of testing they had had two weeks of formal instructions on English and Dutch prosody. Participants were paid an hourly rate of €8 and received a bonus of €10 after attending all sessions of the experiment.

---

[2] One of the participants followed the American English programme. All other students specialised in British English.

4.1.2   Materials

The pre- and post-test had different lists of test items (see Appendix A and B). Each trial

contained three or four sentences: a pre-recorded remark, an optional phrase with background

information about the remark, the sentence the participant was asked to respond with, and

lastly, the reminder that the participant was expected to respond in a "sarcastic manner" in all

cases. An L1 British English speaker recorded the remarks for both the pre- and post-test.

This speaker knew that participants were supposed to respond to the remarks in a sarcastic

manner and was asked to adjust his prosody to provoke sarcastic responses. Some remark-

response pairs were adapted from previous studies on sarcasm (Ackerman, 1983; Kreuz &

Glucksberg, 1989; Capelli et al., 1990; Cheang & Pell, 2008; De Jong, 2014). More items

were created to increase the number of stimuli, so that there were twenty-four experimental

items for the pre-test and twenty-four items for the post-test. The response sentences were of

three sentence types: declaratives, tag-questions, and what-exclamations. Particle sentences

were excluded because a British native speaker involved in the study believed them to be too

American. Additionally, previous studies have already confirmed the hypothesis that these

particle sentences are lexically marked for sarcasm and therefore receive higher ratings on

sarcastic sounding (De Jong, 2014; Kreuz & Caucci, 2007). Responses within each sentence

type were controlled for maximal comparability in length and syntactical complexity. Short

forms were used wherever they would appear in natural speech. Additionally, three L1 British

English speakers unconnected to the current investigation checked the responses for maximal

ambiguousness on both a lexical and a syntactical level, to ensure response sentences did not

have a sarcastic bias. Declaratives were all simple short sentences that contained an adjective

and object noun, or an adjective and adverb. For example:

Remark: I heard Peter, the skinny kid with glasses, is going to beat you up after school.

Background: Your friend knows that you're not afraid of Peter.[3]

Response: *That's very scary.*

Tag-questions were comprised of simple sentences that contained an adjective or adverb and were followed by a negative tag. For example:

Remark: I think I didn't even get one right on that test.

Response: *You did well this time, didn't you?*

What-exclamations all followed the structure [WH indefinite article, Adj N]. For example:

Remark: The arrogant front-runner finished dead last.

Response: *What an amazing result!*

4.1.3 Procedure

The pre- and post-test were individual sessions and were held in sound attenuated booths in the Linguistics department of Utrecht University. The sessions were recorded using a ZOOM 1 digital recorder.

First, participants were given written instructions in English (see Appendix C and D). In both the pre- and post-test, the participant's task was to respond in a sarcastic manner to pre-recorded remarks by a native speaker of British English, following De Jong (2014). They were asked to imagine that they were having short telephone conversations with a British

---

[3] When necessary, test items contained background information about the remark to create a sarcastic bias for the participant's response.

English 'friend' who made remarks about fictional people or events. The instructions furthermore notified participants that they were to make use of prosody to express sarcastic intent, as the sentences they were supposed to respond with were lexically ambiguous. Participants were encouraged to repeat the trial until they were satisfied with their sarcastic sounding. The written instructions for the post-test highlighted what participants had learned in prosody training (see Appendix D). Participant's responses were pre-scribed. The pre-recorded remark, optional background information, the participant's response, and a reminder participants were supposed to respond in a "sarcastic manner" were displayed on-screen using Microsoft PowerPoint. The pre-recorded remark played automatically for each trial and could be repeated by using the mouse to click on a sound icon displayed on the middle of the computer screen. Second, the experimenter asked participants to voice any questions they might have had and to summarise the task to make sure they understood it correctly. Third, the experiment started by completing six practice trials with the experimenter still present in the sound booth. The experimenter left the sound booth for the experimental trials.

Participant's responses were extracted from the recordings using Praat (Boersma & Weenink, 2014). When participants had made multiple attempts at producing the same response, the final attempt was chosen.

4.2 Prosody Training

In week two, all participants who participated in the pre-test were trained in the prosodic markers of sarcasm in English. The training consisted of (1) a presentation highlighting prosodic markers of sarcasm in English, (2) guided practice using audio and visual feedback and (3) individual evaluation on participant's production in small groups. The training took approximately two hours.

Two L1 speakers of British English (1 male, 1 female, M_{age} = 23.5, SD = 2.5), both unconnected to the present investigation, recorded the pre-test materials used in the training. One speaker was a bilingual speaker of British English and Dutch. Another L1 speaker of British English (one of the experimenters) evaluated the L1 recordings on a five-point scale of "not sarcastic" to "very sarcastic". The three highest rated response sentences per sentence type were selected to act as model realisations of sarcasm in English, resulting in nine response sentences to be used in the training. Because the bilingual speaker's recordings received higher overall ratings from one of the experimenters, only her recordings were used for the training.

In the first part of the training, a presentation was given to all participants and their classmates who also attended the course. After explaining basic prosodic concepts, the students were told that sarcasm in English is marked by a wide pitch range, and slower and louder speech (Anolli et al., 2002; Rockwell, 2000). They were told to place strongest emphasis on the most significant content word according to the context (usually the subject, adjective or adverb), with a delayed peak. Explanation was aided by example recordings and pitch contours from the native speakers.

For the second part of the training, all participants received a booklet containing pitch contours and written instructions for the nine response sentences that were recorded by the bilingual speaker (see Appendix E). The words that the bilingual speaker stressed were circled on text-grids that contained orthographic representations of the response sentences. In what-exclamations and tag-questions the boundary tone was accentuated with an arrow placed on the relevant part of the pitch contour. Participants were instructed to listen to the recordings and study the accompanying pitch contour. Subsequently, they had to record the response sentences themselves using Praat (Boersma & Weenink, 2014). The participant's task was to practise their sarcastic production by plotting their own pitch contours and comparing it to the

target pitch contour. Participants were then given the opportunity to practice individually for approximately forty-five minutes.

In the third part of the training, participants joined the experimenters (including an L1 speaker of British English) in groups of two or three for a feedback session. The prosodic production of sarcasm was evaluated in three sentences per participant, one of each sentence type. Their sarcastic production was recorded, played back, and visualised using Praat (Boersma & Weenink, 2014) to make suggestions more clear when necessary.

4.3 Rating Experiment

4.3.1   Participants

Twelve L1 speakers of English (3 males, 9 females, M_{age} = 20.0, SD = 2.2) were recruited as raters at the University of Leeds. Raters did not have any experience of Dutch. Some raters had very limited experience of other, non-Germanic languages. All were monolingual speakers of English that were brought up and educated in the UK and were considered to speak standard British English by the experimenter (who is an L1 speaker of British English). They received £7 for their participation.

4.3.2   Materials

The data that was rated in the rating experiment were the participants' response sentences that were recorded in the pre- and post-test. In total there were 576 sentences to be evaluated. Because rating 576 items would be a very time consuming task for twelve raters, the raters were divided into two groups of six raters. Group A rated the pre-test response sentences of half of the speakers and the post-test response sentences of the other half, meaning each rater evaluated 288 response sentences. Group B rated the items that were not rated by group A.

The response sentences were presented to the rater blocked by speaker. The order of speakers was randomised per rater session, as well as the trials within each speaker. The rating task did not include neutral/sincere fillers.

4.3.3   Procedure

The rating task was set up in ZEP, an experiment control software application (Veenker, 2014). Raters received written instructions (see Appendix F), which told them they were going to listen to several non-native speakers of English who were in a simulated telephone conversation with a friend. They were told that the speakers were trying to sound sarcastic in all their responses. Additionally, they were notified that speakers occasionally mispronounced a word, but that this should not influence their rating. The task was to rate how sarcastic a response sounded on a five-point scale; the left-most circle stood for "not sarcastic" and the right-most circle for "very sarcastic". Raters could listen to each response up to three times by clicking a "listen" button. If they were satisfied with their response, they could click on "next". Additionally, raters were asked to indicate how certain they were of their rating on a second five-point scale that was displayed underneath the first scale. The left-most circle stood for "not certain", the right-most circle for "very certain". Raters started the experiment by completing six practice trials. The data was saved by ZEP for later analysis.

5. Results

In this section the results of the experiments will be presented. The results consist of the sarcasm scores the raters awarded the recordings from the pre- and post-test. An alpha level of .05 was observed in the analysis of all results. Statistical analysis was conducted in R (R

Development Core Team, 2013). A Shapiro-Wilk test revealed that the distribution of average scores was significantly non-normal (W = .98, p < .00), with significant skew and kurtosis. The assumptions of conventional statistical analysis methods are therefore not met, and robust alternatives were used (Wilcox, 2012; Zhang & Yuan, 2013).

5.1 Inter-Rater Agreement

As the twelve raters were divided into two groups that did not evaluate the same material, the agreement between raters was analysed by calculating the robust Cronbach's alpha coefficient per group (Zhang & Yuan, 2013). Rater group A (M = 2.90, SD = .61) showed a robust Cronbach's alpha value of .74. Rater group B (M = 2.87, SD = .61) showed a robust Cronbach's alpha value of .69. These Cronbach's alpha coefficients are both acceptable.

5.2 Mixed-Design Repeated Measures ANOVA

A repeated measures analysis of variance (ANOVA) was performed on trimmed means to look at main effects. The mean was trimmed by 20%, reducing the effect outliers have on average scores. There was one dependent variable, namely the level of sarcasm, and two within-subject independent variables, namely test session and sentence type. See Table 1 on the next page for the descriptive statistics for the variables.

Descriptive Statistics

Sentence type	Test session	N	Mean	Std. Deviation	Std. Error Mean
declarative	pre-test	96	2,807	,944	,096
	post-test	95	3,354	,894	,092
tag-question	pre-test	96	2,483	,737	,075
	post-test	96	2,793	,817	,083
what-exclamation	pre-test	96	2,773	,866	,088
	post-test	96	3,097	,852	,087

Table 1. Descriptive statistics for sentence type x test phase.

5.2.1 Effect of Training

A highly significant main effect of training was found ($Q = 27.23$, $p < .001$). The mean score

for the post-test ($M = 3.08$, $SD = .54$) showed a significant increase ($MD = -.39$) relative to

the mean score for the pre-test ($M = 2.69$, $SD = .44$).

The effect of training was also analysed for each sentence type separately. Post hoc

pairwise comparisons using Wilcox' dependent variation of Yuen's trimmed-means

comparison method revealed significant effects of training for all three sentence types

(Wilcox, 2012). The difference between the pre-test and post-test scores was highly

significant for declaratives, $T_y(57) = -4.47$, $p < .001$, with post-test scores receiving higher

scores than pre-test scores. For what exclamations the comparison between test sessions was

significant, $T_y(57) = -3.31$, $p = .002$, with post-test scores receiving higher scores than pre-test

scores. As was this the case for tag-questions, $T_y(57) = -2.27$, $p = .03$, with post-test scores

receiving higher scores than pre-test scores.

5.2.2 Effect of Sentence Type

The main effect of sentence type was highly significant ($Q = 14.18$, $p < .001$), which means

that the Dutch learners of English received significantly higher scores for some sentence types

than others. These results are depicted in Figure 1 on the next page.

The effect of sentence type was furthermore analysed within each test session. Post

hoc pairwise comparisons using Wilcox' dependent variation of Yuen's trimmed-means

comparison method revealed significant differences (Wilcox, 2012), except for the

comparison between declarative sentences and what exclamations in the pre-test, which

revealed no significant difference, $T_y(57) = -.52$, $p = .61$. In the pre-test the comparison

between declarative sentences and tag questions was significant, $T_y(57) = 2.33$, $p = .02$, with

declaratives receiving higher scores than tag-questions. There was also a significant

difference between what-exclamations and tag-questions in the pre-test, $T_y(57) = 2.22$, p = .03, with what-exclamations receiving higher scores than tag-questions. Within the post-test all comparisons between sentence type revealed significant results. The difference in mean score between declaratives and what-exclamations was significant, $T_y(57) = -2.33$, p = .02, with declaratives receiving higher scores than what-exclamations. Between declaratives and tag-questions a highly significant difference was found, $T_y(57) = 5.57$, p < .001, with declaratives receiving higher scores than tag-questions. Another highly significant result was found in the post-test between what-exclamations and tag-questions, $T_y(57) = 4.13$, p < .001, with what-exclamations receiving higher scores than tag-questions.
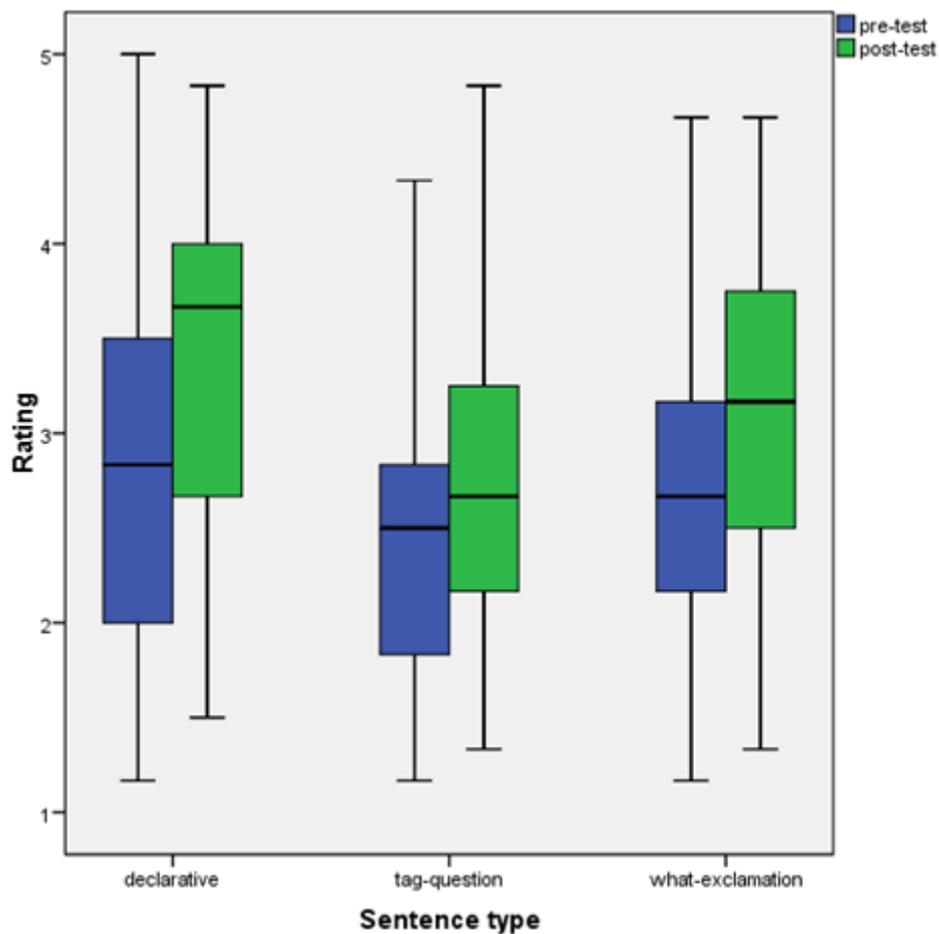

Figure 1. Rating scores per sentence type and test session

5.2.3   Interaction Training x Sentence Type

The interaction between training and sentence type was significant ($Q = 3.77$, $p = .02$), which shows that sarcasm scores did not increase equally for all sentence types across test sessions. The difference between average pre-test score and post-test score is larger in declarative sentences ($MD = -.55$) than in tag questions ($MD = -.31$) and what exclamations ($MD = -.32$).

6.   Discussion and Conclusions

6.1 Training

The main objective of this study was to examine whether Dutch learners of English could improve the accuracy in which they use prosody to convey sarcastic speaker intent after receiving prosody training in the markers of sarcasm in English. Statistical analysis of the results show a highly significant increase in the level of sarcasm scores for the post-test relative to the pre-test, which means the first hypothesis is supported by the results. The effect of training was significant for all sentence types between test sessions.

The results indicate that explicit training in the suprasegmental features of L2 speech is effective. We can thereby conclude that second language education could indeed profit from explicit teaching of the prosodic systems of second languages, something that is not commonly done in formal education at the moment.

6.2 Sentence Type

The effect of sentence type was investigated within test sessions as well as the effect of training within sentence type.

The effect of training within sentence type was significant for all sentence types. However, the interaction found between sentence type and test session indicates that participants improved their scores more in some sentence types than in others, contradicting

our second hypothesis. The analysis showed that participants improved most in declarative sentences. Declaratives received the highest average scores in both the pre- and post-test. Proficiency might influence the effect size of training; Dutch learners might have learned more from the training on declaratives, because they were perhaps more proficient in this structure. However, there is no empirical data to support that statement.

The present study found a highly significant effect of sentence type. Post hoc analysis revealed significant differences between sentence types within test session, except for the difference found between declaratives and what-exclamations in the pre-test. It was previously hypothesised by De Jong (2014) that sarcasm is easier to produce in some structures than in others. She predicted that tag-questions and particle-sentences would receive higher sarcasm scores, because these structures supposedly contain additional non-prosodic markers of sarcasm. Although her results supported her prediction for particle-sentences, this was not the case for her prediction for tag-questions, as those were rated lower than what-exclamations and tag-question scores did not significantly differ from declarative scores. Her surprising results on tag-questions could not be explained by the literature that was reviewed in her paper.

The present study found slightly different effects for sentence type than De Jong (2014) did: declaratives received higher scores than what-exclamations, although this difference was not significant in the pre-test scores, and tag-questions received significantly lower scores than both declaratives and what-exclamations (see Figure 2 in section 5.2.3). The prediction was that declaratives and what-exclamations would receive similar scores, because these two structures lack additional non-prosodic markers of sarcasm. While this was true for the pre-test data, a significant difference was found in the post-test data. More research is necessary to determine what causes the difference between declaratives and what-exclamations. One possible explanation might be the dissimilarity in frequency of these two

structures; what-exclamations are more syntactically marked than declaratives, which may mean that what-exclamations are less frequent in L2 speech than declaratives. Consequently, L2 speakers may be less proficient in the prosodic structure of what-exclamations relative to that of declaratives. Another finding that was not anticipated was the low scores for tag-questions. These structures were hypothesised to be more sensitive to sarcastic interpretation, because tag-questions are often used rhetorically (Kreuz & Caucci, 2007, p. 2). Consequently, we expected scores for tag-questions to be higher than scores for the other sentence types. The literature reviewed in this paper does not provide an explanation for this finding. In sum, the findings on the effect of sentence type need further explanation, especially in regards to the interaction that was found between sentence type and training.

6.3 Limitations

In this section, some of the limitations of the current study are presented.

Firstly, it has to be mentioned that the task in the pre- and post-test did not elicit completely spontaneous speech from the participants, as they were asked to respond with given sentences. Rockwell (2000) proposed that posed sarcasm is easier to perceive than spontaneous sarcasm. Consequently, it is unclear what effect the training will have had on spontaneous production of sarcasm by L2 speakers.

Secondly, some participants may be more proficient in sarcasm in their first language or use sarcasm more frequently in their first language. It is unclear whether proficiency in sarcastic prosody in the L1 affects proficiency in sarcastic prosody in the L2. Previous research suggests that the prosodic markers of sarcasm in Dutch are different from English (De Jong, 2014). It is also possible that participants who frequently use sarcasm in their L1 are more comfortable using such a rhetorical device. Speakers that generally do not use sarcastic speech might feel they are telling lies, which may affect their prosodic realisation.

Third, besides a small data set from the bilingual speaker that participated in the present study, Dutch production of sarcasm was not included in any analysis, so no extensive contrastive analysis could be performed on the markers of sarcasm between English and Dutch. With regards to the training it would be helpful to know what the prosodic markers of sarcasm are in Dutch, so that training can also instruct students on what *not* to do, rather than just presenting the prosodic features of model L2 realisations of sarcasm.

Additionally, no control group participated in the pre- and post-test. There is a small chance the increase in sarcasm scores cannot be attributed to the training, but rather to some other variable that was not taken into account. We tried to reduce this chance by including practice trials and using different test items for the pre- and post-test.

Finally, the sarcasm scores are dependent on both the participants' production and the raters' perception of sarcasm. It is possible that there is a trade-off effect between acoustic and non-acoustic markers, i.e., tag-questions may have a sarcastic bias that may consequently cause speakers to rely on lexical cues to convey their sarcastic intent instead of using prosody. An acoustic analysis of sarcastic productions across the sentence types could determine whether this is the case.

6.4 Future Research

In this section, plans for future research are mentioned, as well as some suggestions for further research.

There are plans to include data from native speakers of British English into the current data set. The pre- and post-tests were also conducted using L1 speakers of British English. In the future, the L1-data will be evaluated by the same raters that evaluated the Dutch speakers' sarcastic production in English. Including the group of native speakers of British English into the data set might give more clarity on the average scores that represent L1 proficiency in the

prosody of sarcasm, so that a comparison can be made between L1 and L2 speakers of English. Consequently, some statements about absolute proficiency may be made.

As existing literature does not provide findings on Dutch prosodic markers of sarcasm, it is necessary to perform acoustic analysis on Dutch sarcastic productions to make further comparisons between English and Dutch sarcasm. Although the findings from the present study, as well as De Jong's (2014) study, suggest that the prosodic markers in Dutch are different from English, empirical evidence is needed to support this hypothesis. The students who participated in the present study may be asked to record Dutch sarcastic sentences in the future. Prosody training in sarcasm may also benefit from learning the Dutch prosodic markers of sarcasm, as training can then specifically address the differences.

Although the present study has found that prosody training has a positive effect on second language learners' prosodic ability, more research is needed to investigate how it could best be implemented in the teaching of second languages. Moreover, more research is needed to explain why some sentence types received higher scores for sarcastic sounding than others.

7. References

Ackerman, B.P. (1983). Form and function in children's understanding of ironic utterances. *Journal of Experimental Child Psychology, 35*, 487-508. Retrieved from http://www.sciencedirect.com/.

Anolli, L., Ciceri, R., & Infantino, M. G. (2002). From "blame by praise" to "praise by blame": Analysis of vocal patterns in ironic communication, *International Journal of Psychology, 37*(5), 266-276, doi: 10.1080/00207590244000106.

Attardo, S. (2000). Irony markers and functions: Towards a goal-oriented theory of irony and its processing. *Rask. International Tidsskrift for Sprog og Kommunication*, *12*, 3-20. Retrieved from http://www.sdu.dk/.

Attardo, S., Eisterhold, J., Hay, J., & Poggy. I. (2003). Multimodal markers of irony and sarcasm. *Humor: International Journal of Humor Research 16* (2), 243–260. Retrieved from EBSCOhost.

Banse, R. & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614-636. Retrieved from http://www.psychologie.uni-bonn.de/.

Beeker, A., Canton, J., Fasoglio, & D., Trimbos, B. (2010). Eindtermen havo/vwo. *Europees Referentiekader Talen*. Nationaal Expertisecentrum leerplanontwikkeling (SLO). Retrieved from http://www.erk.nl/docent/streefniveaus/havo/.

Boersma, P., & Weenink, D. (2014). Praat: Doing Phonetics by Computer [Computer program]. Version 5.3.62. Retrieved from http://www.praat.org/.

Brådvik, B., Dravins, C., Holtås, S., Rosén, I., Rynding, E., & Ingvar, D. H. (1991). Disturbances of speech prosody following right hemisphere infarcts. *Acta Neurologica Scandinavica, 84*(2), 114-126. Doi: 10.1111/j.1600-0404.1991.tb04919.x.

Bryant, G. A., & Tree, J. E. F. (2002). Recognizing Verbal Irony in Spontaneous Speech. *Metaphor and Symbol, 17*(2), 99-119. doi: 10.1207/S15327868MS1702_2.

Capelli, C. A., Nakagawa, N., & Madden, C. M. (1990). How Children Understand Sarcasm: The Role of Context and Intonation. *Child Development, 61*(6), 1824-1841. Retrieved from http://www.jstor.org/.

Carey, M. D., Mannell, R. H., & Dunn, P. K. (2011). Does a rater's familiarity with a candidate's pronunciation effect the rating in oral proficiency interviews? *Language Testing, 28*(2), 201-219. Retrieved from http://ltj.sagepub.com/.

Cheang, H. S., & Pell, M. D. (2008). The Sound of Sarcasm. *Speech Communication, 50*(5), 366-81. doi:10.1016/j.specom.2007.11.003.

Cheang, H. S., & Pell, M. D. (2011). Recognizing Sarcasm Without Language: A Cross-Linguistic Study of English and Cantonese. *Pragmatics & Cognition, 19*(2), 203-23. doi: 10.1075/pc.19.2.02che.

Creusere, M.A. (1999). Theories of Adults' Understanding and Use of Irony and Sarcasm: Applications to and Evidence from Research with Children. *Developmental Review, 19*, 213-262. Retrieved from http://www.idealibrary.com/.

Dews, S., Kaplan, J., & Winner, E. (1995). Why not say it directly? The social functions of irony. *Discourse Processes, 19*(3), 347-367. doi: 10.1080/01638539509544922.

Haiman, J. (1998). Talk is Cheap: Sarcasm, Alienation, and the Evolution of Language. Oxford University Press, Oxford. Retrieved from http://www.jstor.org/.

Irony, n. *OED Online*. December 2014. Oxford University Press. Retrieved 8 January 2015 from http://www.oed.com/.

Jong, D. de. (2014). *Interlingual sarcasm: Prosodic production of sarcasm by Dutch learners of English* (Unpublished bachelor's thesis). Utrecht University.

Jorgensen, J. (1996). The functions of sarcastic irony in speech. *Journal of Pragmatics 26*, 613–634.

Kreuz, R. J. & Caucci, G. M. (2007) Lexical Influences on the Perception of Sarcasm. *Proceedings of the Workshop on Computational Approaches to Figurative Language*,1–4. Retrieved from ACM Digital Library.

Kreuz, R.J., & Glucksberg, S. (1989). How to Be Sarcastic: The Echoic Reminder Theory of Verbal Irony. *Journal of Experimental Psychology: General, 118*(4), 374-386. Retrieved from http://ovidsp.ovid.com/

Laval, V. (2005). French-Speaking Children's Understanding of Sarcasm: The Role of Intonation and Context. *Journal of Speech, Language, and Hearing Research, 48*(3), 610-20. doi:10.1044/1092-4388(2005/042).

Missaglia, F. (2007). Prosodic training for adult Italian learners of German: The Contrastive Method. *Non-Native Prosody: Phonetic Description and Teaching Practice*. Comp. Trouvain, J. & Gut, U.. Mouton de Gruyter, Berlin. 237-52. doi: 10.1515/9783110198751.2.237.

Oring, E. (1994). Humor and the suppression of sentiment. *Humor: International Journal of Humor Research 7*, 7-26.

Pincus, N. (2014). Teaching Intonation With Explicit Instruction and Audio-visual Techniques. *INTED2013 Proceedings*. 7th International Technology, Education and Development Conference, Valencia, Spain. 3428-435. *Iated Digital Library*. Retrieved from http://www.academia.edu/.

R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org.

Rockwell, P. (2000). Lower, Slower, Louder: The vocal cues of sarcasm. *Journal of Psycholinguistic Research, 29*(5), 483-495. Retrieved from http://link.springer.com/.

Sarcasm, n. *OED Online*. December 2014. Oxford University Press. Retrieved 8 January 2015 from http://www.oed.com/.

Veenker, T.J.G. (2014). The Zep Experiment Control Application (Version 1.8) [Computer software]. Beexy Behavioral Experiment Software. Retrieved from http://www.beexy.org/zep/.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye moments. *Cognition, 99*(2), 63-72. doi: http://dx.doi.org/10.1016/j.cognition.2005.07.001.

8. Appendices

A. Pre-test stimuli

Sentence type	Pre-recorded remark	Participant's response	'Common ground' remark
declarative	My aunt smokes a pack a day.	She's a healthy lady.	
declarative	I went for a run and I came back dripping wet.	It's a beautiful day outside.	
declarative	I haven't seen a waiter yet since we were shown a table half an hour ago.	The service's really good here.	
declarative	My plane was an hour late.	Ryanair's always reliable.	
declarative	I bought a new game and I'd thought it would be too hard, but I learned in five minutes.	You're a smart gamer.	You know that your friend is not so good at computer games.
declarative	Tomorrow's class is going to be about plants.	That'll be great fun.	You and your friend aren't interested in plants at all.
declarative	I heard Peter, that skinny kid with glasses, is gonna beat you up after school.	That's very scary.	Your friend knows that you're not afraid of Peter.
declarative	My brother was accepted to the police academy.	Your parents must be proud.	You know that you friend's parents don't agree with his brother's career choice.
what exclamation	My mother-in-law always smirks and snorts loudly when I misspeak	What a respectful gesture!	
what exclamation	The arrogant front-runner finished dead last.	What an amazing result!	
what exclamation	Today, playing football, I slipped and fell and the ball bounced off my head.	What a brilliant header!	
what exclamation	My piano performance has been cancelled.	What a terrible shame!	You know that your friend doesn't feel ready for this performance.
what exclamation	My sister phoned just now and told me that her job interview went very badly.	What a surprising outcome!	You and your friend had foreseen his sister's failure because she didn't prepare for it.
what exclamation	My father had made a five course Christmas dinner; the next day we were all sick with food poisoning.	What an accomplished chef!	
what exclamation	I failed my essay on parliamentary process. I guess I just don't know much about politics.	What a shocking announcement!	You've always known that your friend is not interested in politics, but he didn't want to admit it.
what exclamation	My grandma bought me a yellow sweater for my birthday.	What a lovely colour!	You know that your friend doesn't like yellow
tag question	I traded my cricket bat for a toy truck, but now I find out it's broken.	You got a bargain, didn't you?	
tag question	My brother wanted to help me move and he dropped my grandfather's clock.	He was a big help, wasn't he?	
tag question	I put my homework off for two hours, but then it only took ten minutes.	That took a lot of effort, didn't it?	
tag question	I think I didn't even get one right on that test.	You did well this time, didn't you?	

tag question	My little sister kicked me in the shins.	Your sister's sweet, isn't she?	
tag question	My first football training coach made us run 5 miles; some of the guys threw up.	He's done a beautiful job, hasn't he?	
tag question	I went fishing but didn't catch anything.	They were gracious guests, weren't they?	
tag question	I've joined a running club and go for a run every evening.	You're a nice neighbour, aren't you?	You know that you friend has joined other sport clubs before but given up eventually.
declarative^p	We've been queuing here for over two hours.	Time flies when you're having fun.	
what exclamation^p	My mother makes the best apple pie.	What a domestic goddess!	Your friend's mother has a history of culinary disasters.
tag question^p	You shouldn't buy things on those dodgy websites.	You've certainly never done so, have you?	You know your friend buys stuff on those dodgy websites all the time.
declarative^p	I bought you another pair of socks for Christmas.	I had a desperate need for those.	Your friend knows everyone always gets you socks for Christmas.
what exclamation^p	Do you know if my boss will be there on Saturday?	What a subtle inquiry!	You know your friend feels uncomfortable around his boss and that he tries to avoid him.
tag question^p	Avoiding the main roads will save you a lot of time.	You sure always avoid them, don't you?	Your friend has a reputation for being late and getting lost.

Notes. The last six items, marked $_p$ in the sentence type column, are practice items. These

items were not included in the results.

B. Post-test stimuli

Sentence type	Pre-recorded remark	Participant's response	'Common ground' remark
declarative	I got one of the lowest grades in the maths test.	That's never happened before.	Your friend often gets the lowest grade in maths tests.
declarative	I heard that the camping trip has been cancelled.	You must be really disappointed.	You know your friend hates camping.
declarative	My uncle keeps telling that stupid joke over and over again.	That joke's hilarious.	
declarative	My baby sister fell asleep on her dinner plate.	That sounds comfortable.	
declarative	My dance partner keeps stepping on my toes.	She's a graceful dance partner.	
declarative	Only three people showed up to my housemate's party last night.	That sounds wild.	
declarative	How's being stuck at home after your accident?	I'm having a great time.	Your friend knows you hate being cooped up inside.
declarative	Did you see my art coursework?	You could be a professional artist.	You know art isn't your friend's forte.
what exclamation	My cat left me a present on the doormat.	What a nice surprise!	
what exclamation	I ordered soup for lunch and found a hair in it.	What a tasty lunch!	
what exclamation	I can touch my nose with my elbow.	What a useful skill!	
what exclamation	My sister's boyfriend ran out of the haunted house screaming like a little girl.	What a brave man!	
what exclamation	I got a lift from Mary. She kept indicating the wrong way at roundabouts.	What a great driver!	
what exclamation	People left the cinema halfway through the film.	What a gripping film!	
what exclamation	My brother's friend came to our house just because he wanted to play our new game.	What a considerate friend!	
what exclamation	I almost fell asleep in class today.	What an engaging lecture!	
tag question	I warned Mark about washing his whites and coloureds together.	He takes advice well, doesn't he?	You and your friend know that Mark hardly takes advice from anyone.
tag question	The weather forecast didn't say it would rain today.	The weather forecast is always right, isn't it?	
tag question	My nephew showed his brand new iphone to everyone at the party.	He's modest, isn't he?	
tag question	I had a busy morning, walking the dog and doing the dishes	You've worked hard, haven't you?	
tag question	Kim turned up at my party even though she wasn't invited.	You were pleased to see her, weren't you?	You know your friend doesn't like Kim.
tag question	My father helped me paint my flat; there's paint splatters everywhere.	He's done a beautiful job, hasn't he?	
tag question	I had to serve all my aunts and uncles drinks all afternoon.	They were gracious guests, weren't they?	You know your friend's aunts and uncles like ordering people around.
tag question	I caught my cat eating the fish from next door's pond.	You're a nice neighbour, aren't you?	You know your friend always lets his cat eat his neighbour's fish.
declarative[p]	We've been queuing here for over two hours.	Time flies when you're having fun.	

what exclamation[p]	I always have to pretend that I like my mother's apple pie.	What a domestic goddess!	Your friend's mother has a history of culinary disasters.
tag question[p]	You shouldn't buy things on those dodgy websites.	You've certainly never done so, have you?	You know your friend buys stuff on those dodgy websites all the time.
declarative[p]	I bought you another pair of socks for Christmas.	I had a desperate need for those.	Your friend knows everyone always gets you socks for Christmas.
what exclamation[p]	Do you know if my boss will be there on Saturday?	What a subtle inquiry!	You know your friend feels uncomfortable around his boss and that he tries to avoid him.
tag question[p]	Avoiding the main roads will save you a lot of time.	You sure always avoid them, don't you?	Your friend has a reputation for being late and getting lost.

Notes. The last six items, marked [p] in the sentence type column, are practice items. These

were not included in the results.

C. Instructions pre-test

Instructions

First of all, thank you for participating in this experiment!

What is your task?
Imagine that you're on the phone with a good friend. You've been chatting for a bit and then your friend makes a remark.

The task is to respond to your friend's remark with a sentence that is presented on a PowerPoint slide. Your response is supposed to *sound* sarcastic. Please note that the sentences are lexically ambiguous. You can give them a sarcastic sounding via prosody.

How does the experiment work?
You will see three (or four) sentences appear on your computer screen:

Your friend's remark
(If necessary, supposedly shared knowledge related to the remark is given)

The response you will give to your friend
(The manner in which you respond, i.e. sarcastic)

Take your time to familiarise yourself with your friend's remark and your response. You can

then listen to your friend's remark by clicking on a speaker icon 🔊 .
You can listen to it as many times as you want.

Next, you utter your sarcastic response. This is also stated in parenthesis below each response. You can have as many attempts as you wish.

If you're satisfied, press the 'enter' key on your keyboard to go to the next slide.

We will start with some practice sentences to familiarise you with the task.

What else do you need to pay attention to?
Please utter your response sentence exactly as it is displayed on your screen. So don't add or omit any words.

Lastly
If you have any questions, please ask the experimenter. You can ask questions during the experiment as well. In the latter case, leave the studio and find the experimenter, who is waiting for you outside the studio.

Have fun!

D. Instructions post-test

Instructions – Part A

The experiment consists of two parts, A and B. In part A of the experiment, your task is identical to that in the experiment in which you participated over a week ago. Very briefly, you are supposed to give a sarcastic-sounding response to a friend's remark on each trial.

Before you get on with the task, please recall the tips you have received previously to sound convincingly sarcastic in British English:
- Talk (much) more slowly than you normally do
- Use a much wider pitch span than you normally do
- Emphasise the content word that is most important in each sentence according to the context with a rise-fall pattern by doing the following:
    o Use a delayed peak on the stressed syllable of the emphasised word so that you rise from a lower pitch to a peak and then fall again in your pitch, instead of falling from a peak directly to a low pitch.
    o In other words, postpone the fall till the second half of the stressed syllable (as in 'big'), or the beginning of the following unstressed syllable (as in 'really').
- Emphasise other content words with a falling pitch pattern (with a somewhat reduced pitch span) if you want to
- Avoid going up in pitch at the very end of a sentence

For your convenience, the instructions used in the previous experiment are repeated here.

What is your task?
Imagine that you're on the phone with a good friend. You've been chatting for a bit and then your friend makes a remark.
The task is to respond to your friend's remark with a sentence that is presented on a PowerPoint slide. Your response is supposed to *sound* sarcastic. Please note that the sentences are lexically ambiguous. You need to give them a sarcastic sounding via prosody.

How does the experiment work?
You will see three (or four) sentences appear on your computer screen:

Your friend's remark
🔊

(If necessary, supposedly shared knowledge between you and your friend related to the remark is given here)

The response you are supposed to give to your friend
(The manner in which you are supposed to respond, i.e. sarcastic)

Take your time to familiarise yourself with your friend's remark and your response.

Next, you utter your response in a sarcastic intonation. This is also stated in parentheses below each response. You can have as many attempts as you wish.

If you're satisfied with your utterance, press the 'enter' key on your keyboard to go to the next slide.

We will again start with some practice sentences to remind you of the task.

What else do you need to pay attention to?
Please utter your response sentence exactly as it is displayed on your screen. Please don't add or omit any words. If this happens, or you misspeak, please utter the sentence again. If you have any questions, please ask the experimenter.

E. Training Document

<u>Training</u>

Last week, all of you were recorded while you were doing a task that asked you to respond sarcastically in English to a hypothetical friend with a given remark. The sentences that you were asked to utter sarcastically were of three types: (1) what exclamations, (2) tag-questions, and (3) declaratives. These sentences were lexically ambiguous, which means you can also use them sincerely.

Based on what we've found in literature and in two native speaker we've recorded, we will provide training on how to sound sarcastic in these three sentence types.

In the training, you will first listen to examples of sarcastic utterances produced by native speakers of British English and be told about the prosodic features involved. You will then be asked to mimic the prosody in these utterances and record your production via the headphone set in Praat. The experimenter and a native speaker of British English will give you feedback on your prosody by comparing your pitch contour with the pitch contour of the example sentence.

Make a recording with Praat

You can make a recording with Praat by clicking on 'New' in Praat Objects. The sample frequency is set at 44100 Hz by default, you don't have to change this.

Click on 'record' when you're ready to record yourself. You can then save the recording to your Praat Objects list by clicking on 'Save to list'. In the Praat Objects window you can select it and see the pitch contour by clicking on 'View and edit'.

**Compare your speech to the native utterance**

After you've recorded yourself, you can compare your sarcastic utterance with the native sarcastic utterance. Check whether you've (1) emphasised the right word, (2) have a wide pitch range (maximum pitch – minimum pitch), (3) use the right final pitch (final rise vs. final fall), (4) speak loudly, and (5) speak slowly.
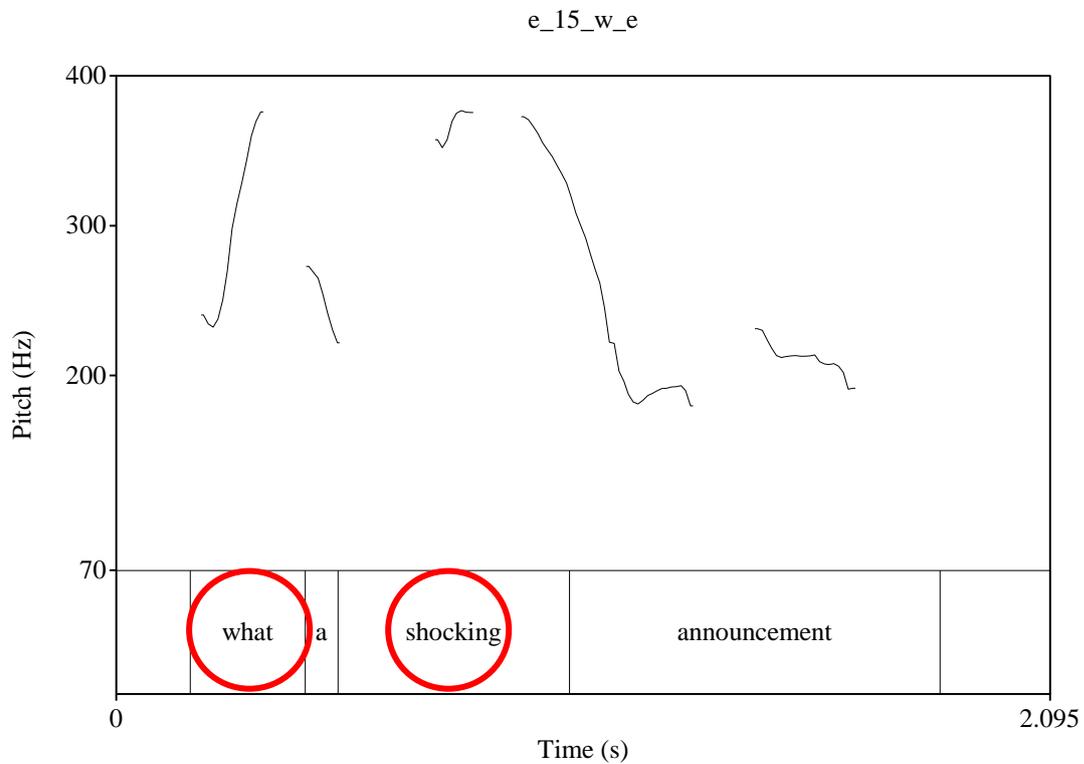
1. **What exclamations**

To sound sarcastic in what-exclamations you:
- Emphasize the subject ("what") and/or the adjective/adverb/quantifier.
- Speak more slowly
- Widen your pitch range.
- Speak loudly.

Practice with the following sentences:
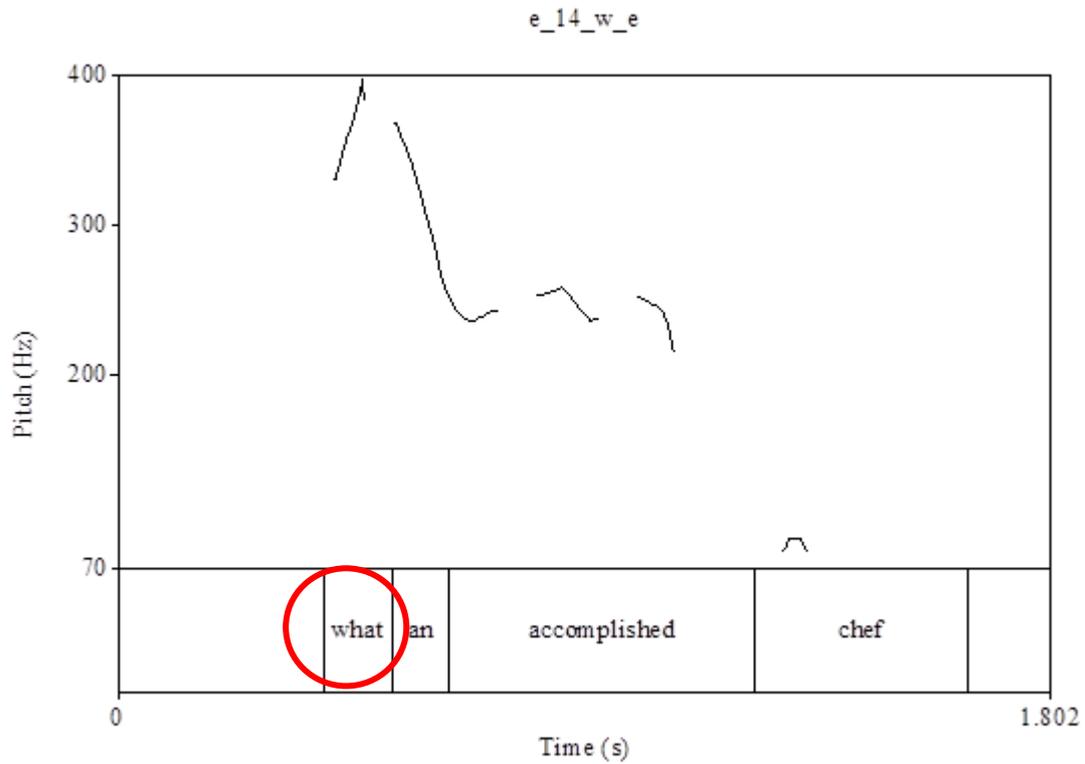
In the pitch contour of the sentence above we see that the subject, "what", was emphasised by a salient rise-falling pitch movement. The pitch range is very wide; there is almost a 200 Hz difference between the emphasised word and the rest of the sentence.

e_16_w_e
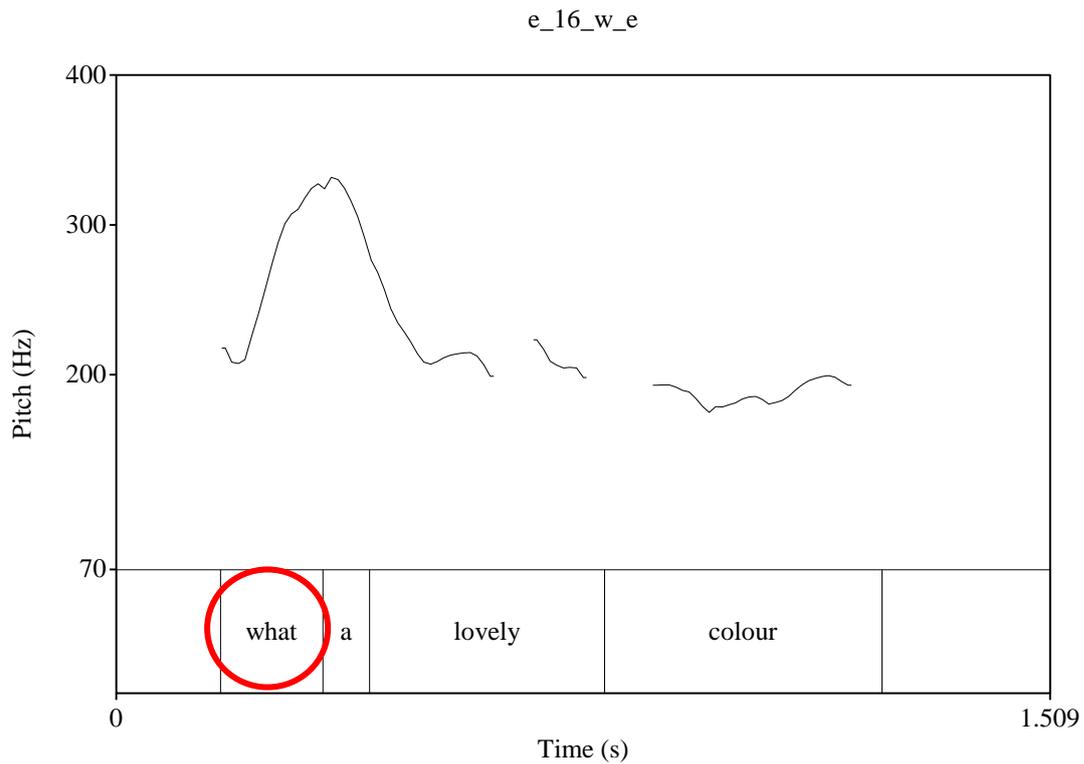
Below there is also an example of a male native speaker. The pitch register of males is lower, which means that their pitch range in Hertz can be smaller; from 200 to 400 Hz is 12 semitones, but from 100 to 200 Hz is also 12 semitones. In this example the minimum pitch is 70 Hz and the maximum pitch 170 Hz. This is a difference of 100 Hz, or 15 semitones.

o_20_w_e
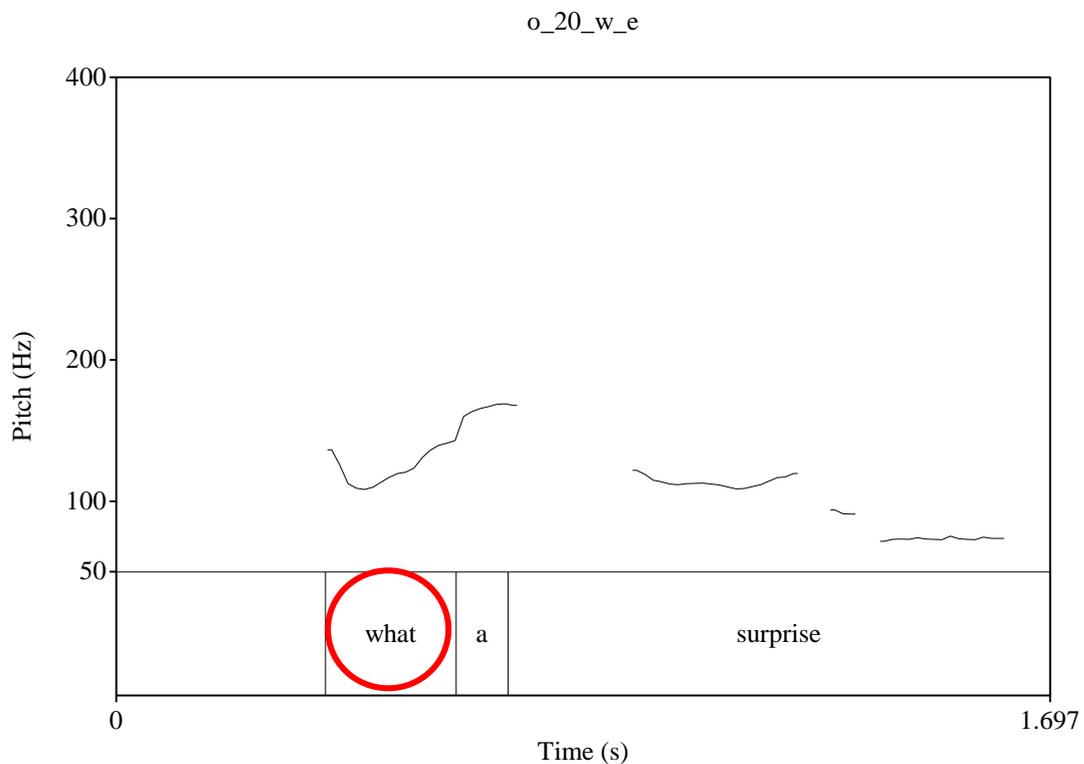
## 2. Tag questions

To sound sarcastic in what-exclamations you:

- Emphasize the subject and the object noun or the adjective or noun in the predicate (e.g. 'sweet' in 'your sister's sweet', and 'shape' in 'in shape'). The pitch movement on the second is less salient than the pitch movement on the former.
- Produce the tag question itself with a fall or a small rise.
- Speak more slowly.
- Widen your pitch range.
- Speak loudly.

Practice with the following sentences:

e_19_t_e
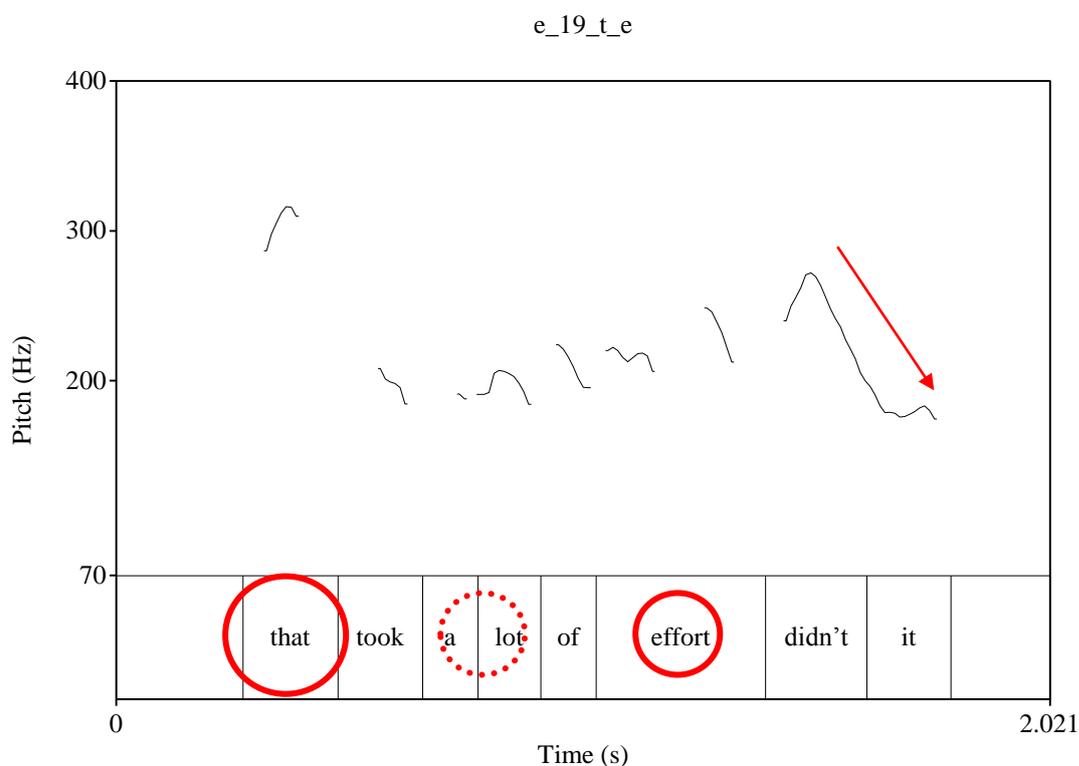
In the sentence above, you could also choose to emphasize the quantifier 'a lot'.

e_21_t_e

e_22_t_e
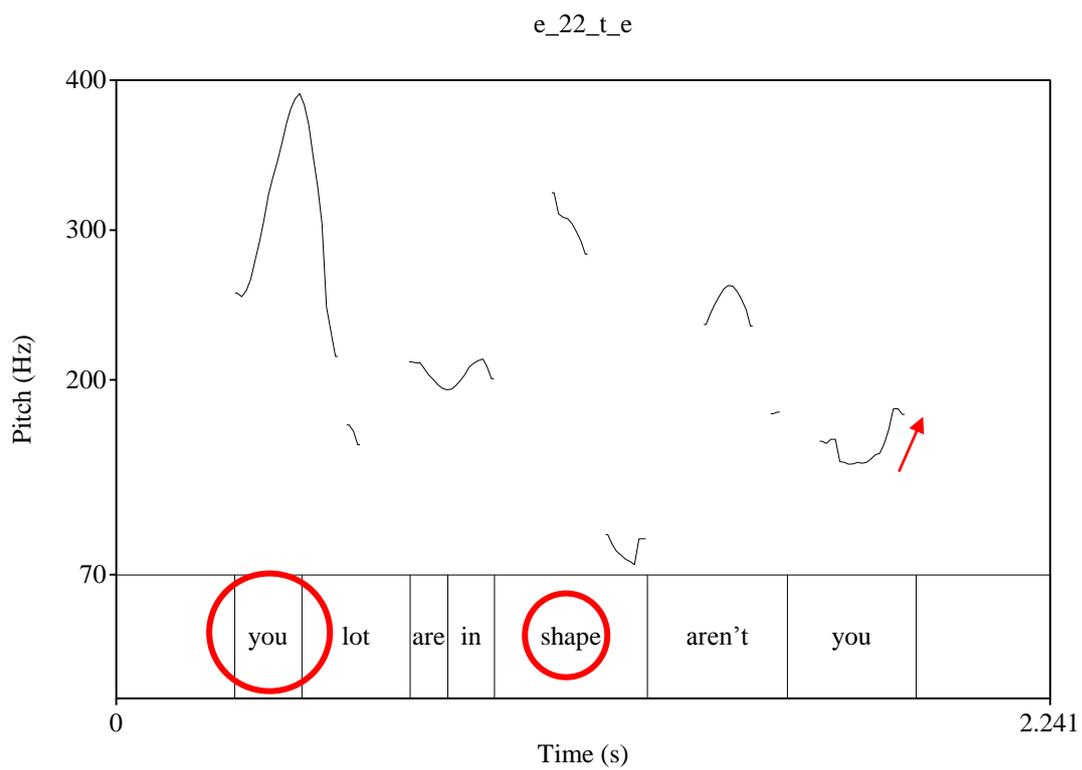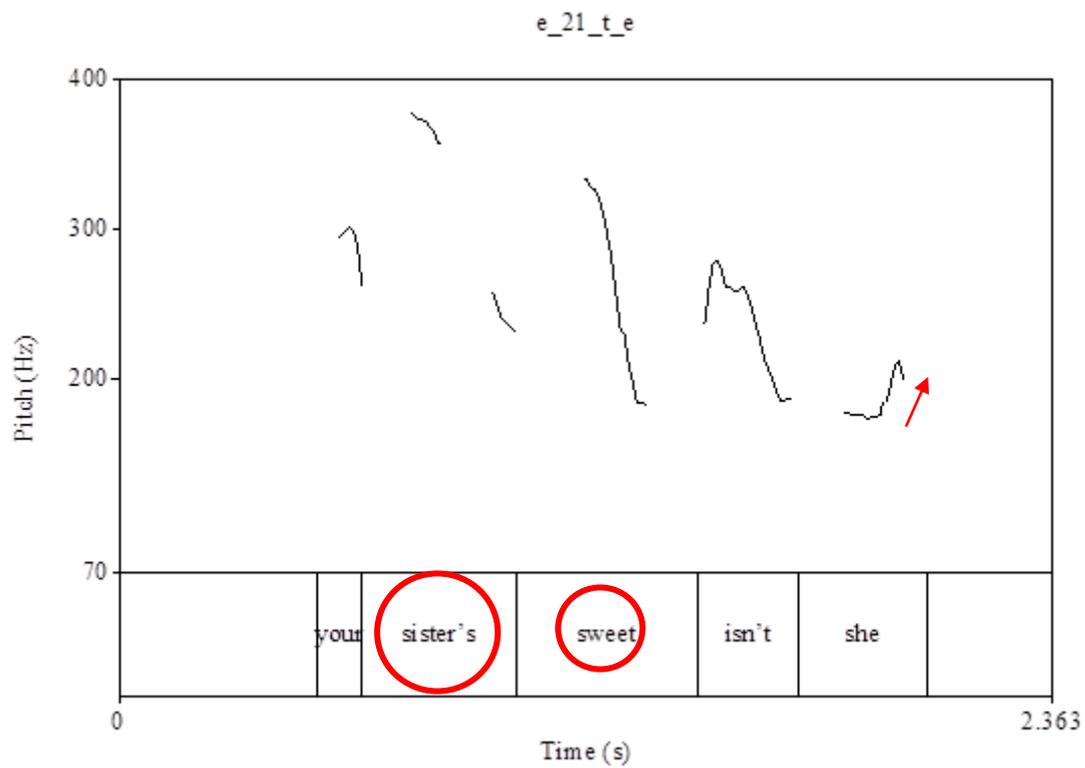
3. Declaratives

To sound sarcastic in what-exclamations you:
- Emphasize the subject and the adjective/adverb/quantifier in the predicate. The pitch movement in the latter is less salient than in the former.
- End the sentence with a fall or a slight rise.
- Speak more slowly.
- Widen your pitch range.
- Speak loudly.

Practice with the following sentences:

e_1_d_e

e_3_d_e

the | service | is | really | good | here

e_8_d_e

your | parents | must | be | proud
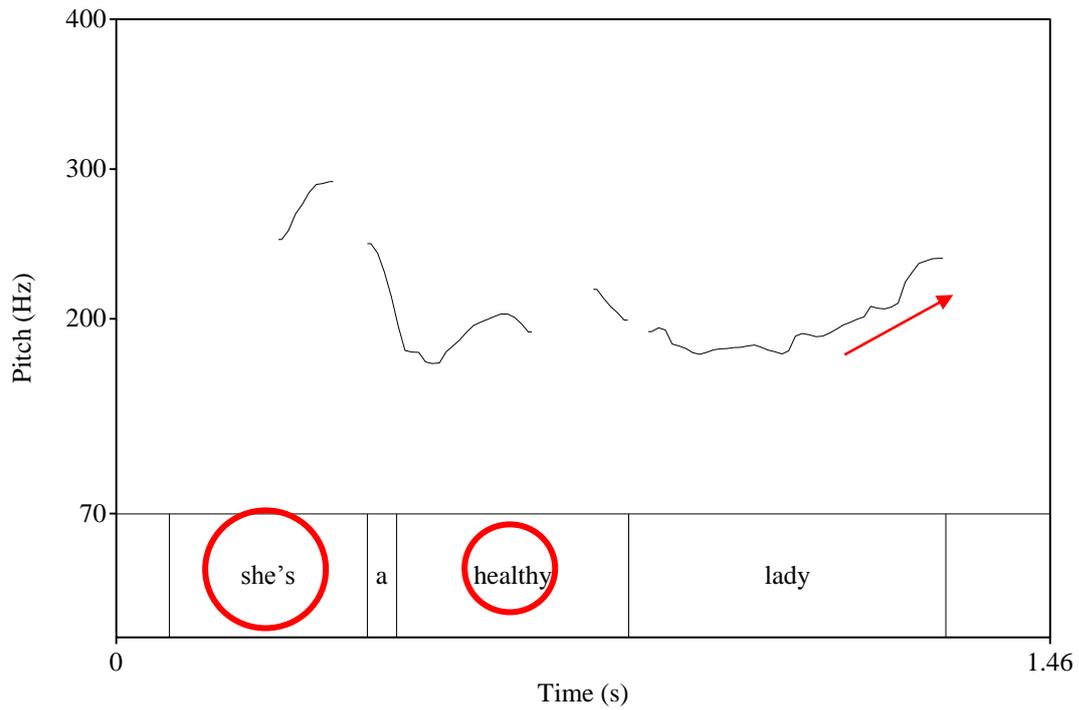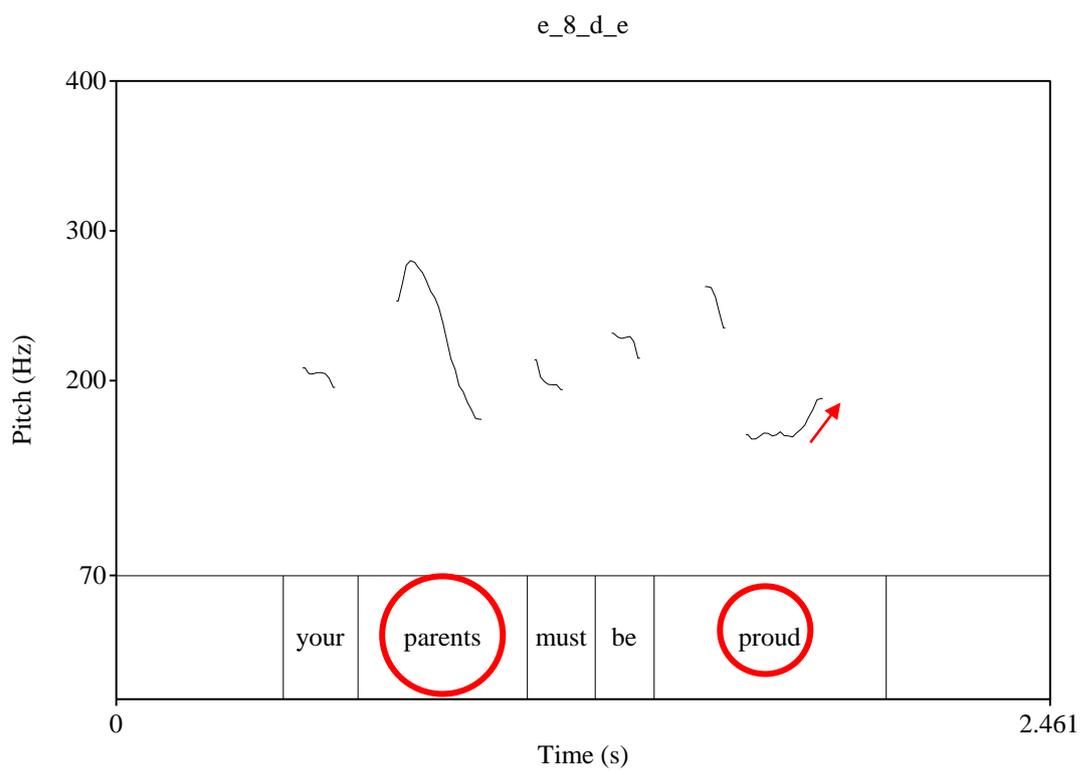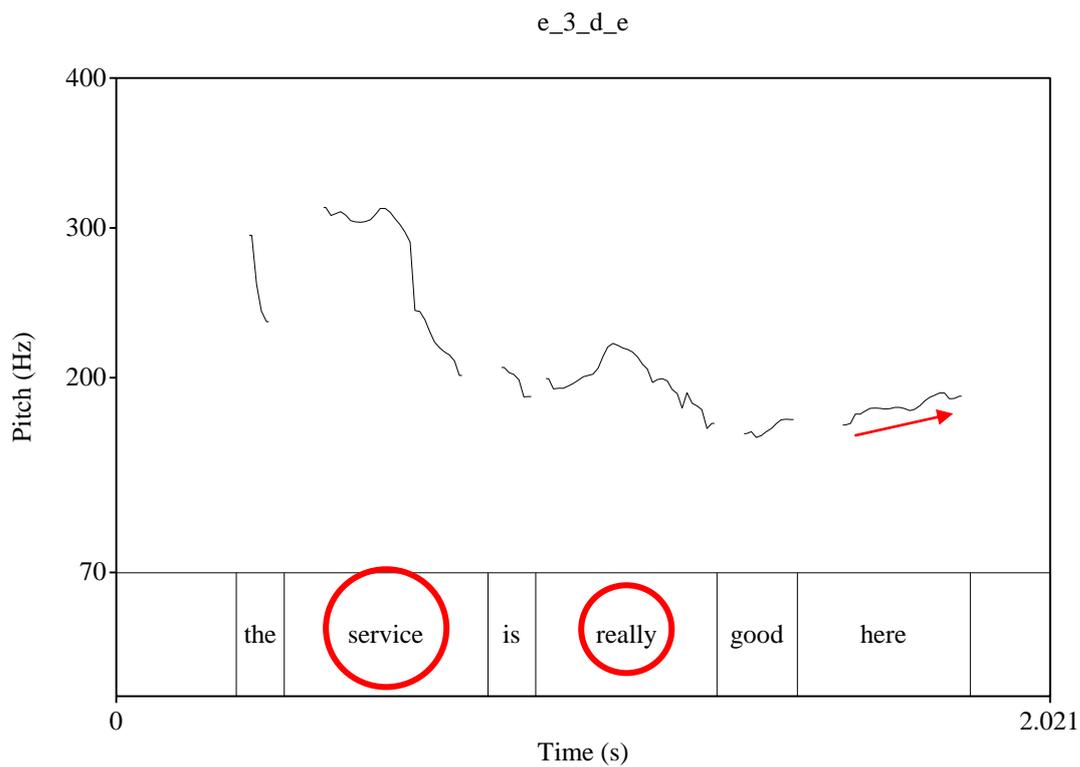
F. Instructions Rating experiment

Instructions

In this research, we would like to know how well non-native speakers of English can use intonation (also known as prosody or melody of speech) to express sarcasm in daily communication.

What you will hear
You will hear a number of sentences spoken by different people following a script. They were replying to remarks made by their friend in simulated telephone conversations. They were aiming to respond in a sarcastic tone in all their responses. However, they might manage better in some conversations than in other conversations; some speakers might be more successful in this than other speakers.

Your task
Your task is two-fold. First, you rate how sarcastic each response sounds to YOU on a 5-point scale. The left-most circle on the scale stands for 'not sarcastic'; the right-most circle stands for 'very sarcastic'. Second, you rate how certain you are of your rating on sarcasm on another 5-point scale. The left-most circle on the scale stands for 'not certain'; the right-most circle stands for 'very certain'. You give your ratings by ticking a circle on the scales on a computer screen.

Not sarcastic   ◯   ◯   ◯   ◯   ◯   Very sarcastic

Not certain   ◯   ◯   ◯   ◯   ◯   Very certain

You can listen to each sentence up to three times. To listen again, click "LISTEN".
When you are done with both scales, click 'NEXT' to move on to the next short conversation.

Structure of the experiment
The experiment starts with 6 practice trials. The experiment proper consists of 288 sentences, 24 per speaker. You proceed at your own pace. You can take short breaks during the experiment if you need to.

Incidental mispronunciation
Please note that occasionally a speaker mispronounced a word. This should, however, not influence your rating of the sarcastic tone.

Enjoy!