

Universiteit Utrecht

Department of Modern Languages

Bachelor's Thesis

Interlingual Sarcasm:  
Prosodic Production of Sarcasm by Dutch Learners  
of English

Name:	Diantha de Jong
Student Number:	3769615
Address:	Graaf van Bloisstraat 56, 2805 RM Gouda
Bachelor:	English Language and Culture
Supervisor:	Aoju Chen
Second Reader:	Koen Sebregts
Course:	Principles and Parameters of SLA (level 3)
Model:	British English
Date:	25 February 2014
Word Count:	7938 words

## **Abstract**

Whether second language learners are able to produce sarcasm effectively in their second language is a question which has received little attention from academics so far. Furthermore, the role of intonation in this context is unclear. The present study was conducted to shed light on the question whether L2-learners are able to express sarcasm using intonation which is recognisable to native speakers. To this end, sarcastic responses of Dutch learners of English of intermediate and advanced L2 proficiency levels were recorded and rated on degree of sarcasm by Dutch and British judges. The results of analyses indicated that Dutch learners of English possibly retain a Dutch intonation of sarcasm in English. Furthermore, the results suggest that sentences containing particle discourse markers are easier to express in a sarcastic manner than other types of sentences.

## **I. Introduction**

Intonation has long been argued to be important in communication. Previous studies have investigated the influence of intonation on the perception of different emotions and its ability to convey communicative information (Morton & Trehub, 2001). We are interested in the role of intonation in expressing sarcasm, and in particular, the ability of second language learners to do so. An extensive body of literature has been published on children's and adults' perception of irony and sarcasm, and their relative use of context and intonation in recognising these meaning attributes in their first language (cf. Capelli, Nakagawa, & Madden, 1999; Creusere, 1999; Creusere, 2000; Katz, Blasko & Kazmerski, 2004; Kreuz, 2000; Laval & Bert-Erboul, 2005; Nakassis & Snedeker, 2002; Recchia, Howe, Ross & Alexander, 2010). However, little attention has been paid to the ability of perceiving and producing irony and sarcasm by L2-learners. Against this background, we investigate how

effectively Dutch learners of (British) English are able to produce sarcasm in English. We are also interested in the role of overall proficiency in this context.

### *Defining Sarcasm*

In the field of communication research, sarcasm is often discussed together with irony.

Various definitions of irony and sarcasm have been put forward in the literature. Therefore, it is necessary to distinguish between different definitions and determine a working definition of sarcasm in the current study.

The general consensus among researchers is that irony contains a difference between the literal meaning of words and the intended meaning of the producer (cf. Bowes & Katz, 2011: 216; Capelli, Nakagawa & Madden, 1999: 1824; Cheang & Pell, 2008: 366; Creusere, 1999: 215; Kreuz & Glucksberg, 1989: 374; Recchia et al., 2010: 255). It is not surprising that the term is defined in such broad terms, because irony appears in a great variety of forms; written form, oral form, and visual form (Rockwell, 2001). Verbal irony is a more specific form of irony, which conveys attitudes “toward some object, event, or person” (Kreuz & Glucksberg, 1989, 374) by using non-literal language.

Sarcasm and irony are often used to refer to the same rhetorical device (Kreuz, 2000, 100). However, sarcasm is not necessarily ironic. Many academics agree that sarcasm differentiates itself from irony by a negative intent of causing pain (Capelli et al., 1999, 1824; Cheang & Pell, 2008, 366; Creusere, 1999, 215; Kreuz & Glucksberg, 1989, 374; Lee & Katz, 1998; Muecke, 1980). However, its connection to irony remains arguable. In an attempt to find distinguishable features for sarcasm, Kreuz and Glucksberg (1989) conducted experiments to investigate the importance of a victim in the recognition of sarcasm. For this purpose they manipulated story dialogues to create multiple versions in which the victim of the story had different degrees of expertise. The experiment also included a control dialogue,

which contained no victim (378). The victim made a faulty prediction, which was commented on. These dialogues were then rated on their degree of sarcasm and irony. The results of this study suggest that when the literal meaning of a speaker was obviously misplaced in the context, no victim was needed for the remark to be judged as sarcastic. Furthermore, the expertise of the victim was of no influence. However, ratings for sarcasm were higher when a victim was involved (382). The study suggests that statements have to be counterfactual in order to raise suspicion of a sarcastic form, but the presence of a victim strengthens this conviction. However, in a more recent study, Kreuz (2000) argues that when a statement is counterfactual, other conclusions can also be drawn, for example, the speaker is lying (100). In an attempt to add to the study of Kreuz and Glucksberg (1989), Lee and Katz (1998) aim to clarify the distinction between sarcasm and irony by investigating the influence of ridicule on ratings of both. Their method was to use passages from the study by Kreuz and Glucksberg (1989) and manipulate these to form two pairs of stories: one with correct and incorrect predictions and the other with ridicule of the other and ridicule of the self (3). They concluded that passages with victims who made incorrect predictions, i.e. who could be ridiculed, received higher ratings of sarcasm. This effect was only true for negative statements. For irony, the researchers found no effect of the presence of a victim to ridicule (5). These findings were supported by the results of a second experiment, in which the degree of ridicule was rated for the same stories as the ones in the first experiment. The stories which received high ratings of sarcasm in the first experiment were also given high ratings for the degree of ridicule they contained (8). The authors suggest that a high degree of ridicule is associated with sarcasm. A criticism of the work of Lee and Katz (1998) is that, unlike in the experiment of Kreuz & Glucksberg (1989), they did not ask participants for their definitions of sarcasm and irony, nor did the researchers present participants with their own explanations of the terms. As a result, participants may have had different notions of these terms, which

potentially make the results hard to interpret. Another argument which requires further investigation is whether the person of ridicule can also be a group or even an object or a viewpoint. In that case, a sense of sarcasm would be raised even if there is no clear victim of the remark, as has been suggested by Kreuz and Glucksberg (1989, 382).

Possibly as a result of the undetermined debate on the distinction between irony and sarcasm, many academics prefer to refer to sarcasm as a subcategory of verbal irony. The term “ironic sarcasm,” occasionally called sarcastic irony, has been coined to refer to a range of common remarks that express the opposite of what is meant in an ironic manner in order to convey a negative or hurtful attitude towards a specific person or topic (Capelli et al., 1990, 1824; Kreuz & Glucksberg, 1989, 374).

As has become clear, there is little convincing evidence that sarcasm can be set apart from irony entirely. For this reason, this paper will use the term sarcasm to refer to the meaning of ironic sarcasm, suggested by Capelli et al. (1990) and Kreuz and Glucksberg (1989).

### *Context and Intonation*

It has been mentioned above that irony can appear in different forms. In order to be recognised in different contexts, sarcasm possesses a particular quality. More specifically, it contains a number of cues to trigger perception, but not all cues are needed for recognition of a sarcastic construction (Capelli et al., 1990, 1825). Context and intonation have been named as two categories of cues that are used in sarcasm perception (Capelli et al., 1990; Creusere, 1999). Many studies regard context as the most important indication of sarcasm. Consequently, a large number of experiments are text-based (Rockwell, 2000, 486). Sentence structure and choice of word may also play a role. According to Kreuz and Caucci (2007), certain structures or words are more sensitive to non-literal interpretation than others (2).

They name tag questions, which are often rhetorical, particle discourse markers which express emotions, such as “gee” and “gosh” (2-3), and repetitions (4) as cues for sarcasm.

Gibbs and O’Brien (1991) review previous psycholinguistic research and argue that adults do not need intonational information to perceive sarcasm, as they are able to deduce verbal irony from contextual discrepancy (529). However, they state that intonation is responsible for a quicker understanding of sarcastic intent (526). In addition, in an experiment conducted by Capelli et al. (1990) it is discovered that children who were only provided with contextual cues were unable to perceive sarcasm more than half of the time, whereas adults attained much higher scores (1836). However, when a sarcastic intonation was used in the stimuli, all children of eight to twelve years old were able to significantly recognise sarcasm (1836). Adults received similar scores as in the cases where intonation was not available, as they were also able to rely on contextual cues. As an explanation, the authors suggest that a “negative tone of voice” and the presence of an object would alert children to sarcasm, but they would not be aware that the speech was non-literal, thereby not perceiving irony, except in the case of the sarcastic form where the intended meaning is the opposite of the literal meaning (1837). For adults, intonation seems to play a less important role, because adults were also able to perceive irony by context alone. This is in accordance with the review by Gibbs and O’Brien (1991). Even adults may misinterpret the subtle use of sarcasm and make errors (Creusere, 2000, 29). However, there seems a definite increase in the ability to perceive sarcasm from childhood and adulthood (Capelli et al., 1990), because adults have acquired the ability to interpret contextual cues in an ironic statement.

However, the importance of intonation should not be underestimated. In an experiment by Morton and Trehub (2001) children from four to ten years old were able to judge the emotion of speakers accurately when the context was not available, in foreign and low-pass filtered speech. However, when context was available, the children under 10 responded

primarily to the context, instead of the intonation, which conflicted with the context half of the time; they described an event with a good ending in sad intonation and an event with a bad ending in happy intonation (841). Morton and Trehub (2001) state that the importance of context in the eyes of children may be a reason why children have difficulty understanding sarcasm, which contains a conflicting context which may be difficult to interpret (841). However, the previously described study by Capelli et al. (1990) suggests adults do not have this problem. Furthermore, intonation of sarcasm may behave differently than intonation of basic emotions such as happiness and anger, which are used in the study by Morton and Trehub (2001). It is possible that intonation of sarcasm cannot be recognised in foreign speech in this way, either because contextual cues are needed or because the intonation is different in foreign languages.

Although there is mixed evidence on the relative importance of intonation compared to context in expressing and perceiving sarcasm, there seems to be some consensus that adults are better able to express and perceive sarcasm than children, and the relevance of intonation is well recognised.

### *Interlingual Sarcasm*

The relevance of understanding another speaker's intent in a multilingual environment is evident in a globalising world, where intercultural interactions are increasingly more common (Cheang & Pell, 2013, 18-19). Intonation may be an important tool for recognising sarcasm in such situation. However, it has been mentioned that the role of intonation in the perception of sarcasm is still unclear. Some researchers argue the existence of an "ironic tone of voice" (Creusere, 1999, 231; Nakassis & Snedeker, 2002). Research by several authors reveals slower speech, greater intensity, nasalisation, and intonational features such as monotony, but

also extreme pitch levels (Attardo & Eisterhold & Hay & Poggi, 2003, 493; Rockwell, 2000, 2007) to be cues for sarcasm use across languages.

However, very recent research into the perception of sarcasm in a foreign language by Cheang and Pell (2013) suggest that the intonation of sarcasm differs in languages, even if the intonational cues of these languages appear similar in theory. Cheang and Pell (2013) state that, although earlier research has established that basic emotions can be recognised by use of intonation in a foreign language, this does not seem to be the case for sarcasm (18, 25). In their experiment, English and Cantonese listeners were presented with sarcastic speech in their own and the other language. Participants were then asked to identify the manner in which each utterance was produced. They were able to choose from sarcastic, sincere, neutral, and humorous manner. The results indicate that, although quite accurate in perceiving the correct intonation manner in their native language, participants were less successful in the foreign language (28). Both English and Cantonese listeners often rated sarcasm in the other language as sincere (25). Furthermore, utterances in the category of humour in Cantonese were often identified as sarcasm by English listeners. The results that speakers were significantly accurate in identifying all the intonation manners in their own language suggest that sarcastic, sincere, neutral, and humorous intonation can be perceived without explicit contextual cues. In addition, a study by Rockwell (2000) suggests that listeners are able to discriminate posed sarcasm from non-sarcasm when context is not available, suggesting the important role of intonation in expressing and perceiving sarcasm. However, the difficulty which the listeners of the study by Cheang and Pell (2013) have in identifying sarcastic intonation in the other language of the experiment strongly suggests that the intonation of sarcasm is not universal.

As a result, the question is raised whether every language has a different and specific intonation of sarcasm. If this is the case, it is possible that L2-learners have to adjust their



intonation in order to convey sarcastic intent in the L2. However, the ability of sarcasm production in the L2 is unclear, as there has been little research on this topic. Cheang and Pell (2013)'s research into cross-lingual perception of intonational cues is a pioneering study in this area. It is possible that sarcasm development in L2-learning is similar to the development of sarcasm use in L1-acquisition. In that case, development of sarcasm production would increase as the learner encounters instances of sarcasm more often and L2-proficiency increases. However, another possibility is that L2-learners retain their native intonation of sarcasm in the L2. For the present study this could mean that Dutch learners of English will not be able to express sarcasm in an intonation which English speakers will find sarcastic.

#### *Other Factors*

Beside context and intonation, there may be other cues of sarcasm. In order to complete the theoretical background of this study, it is necessary to briefly analyse previous research on other influencing factors.

Rockwell (2001) analysed facial expressions of participants who were placed in a setting in which they were able to produce instances of spontaneous sarcasm. The utterances were rated on the degree of sarcasm and the facial expressions corresponding with sentences with the highest and lowest scores of sarcasm were used for analysis (49). The use of eyebrow movements had been argued to be a cue for sarcasm use. However, Rockwell's research showed only significant difference in mouth movement between sarcastic and non-sarcastic utterances (49). Attardo et al. (2003) investigated the use of facial expressions in combination with sarcasm in television shows. They state that previous research indicate raised or lowered eyebrows, squinting or rolling eyes, nodding, smiling, and winking as visual cues for sarcasm (245). In addition, they argue that another expression indicates sarcasm, namely a "blank face", an emotionless facial expression (254). However, as the authors indicate, by using

acted utterances from television shows, facial cues of sarcasm may have been exaggerated in order to make it clear to the audience that sarcasm was being used (246). Whether these expressions are retained in such clarity in spontaneous conversation remains unclear.

However, the previously described studies indicate that there are more markers of sarcasm besides context and intonation which should be taken into account when investigating sarcasm.

### *Research Questions and Hypotheses*

The current study aims to investigate sarcasm production by L2-learners of English and a possible relation between effectiveness of sarcasm production and L2-proficiency level.

Furthermore, different types of sentences are used in order to determine which constructions are more difficult to utter in sarcastic intonation. We will use the following research questions:

*RQ<sub>1</sub>: Are Dutch learners of English able to express accurate intonation of sarcasm in English?*

*RQ<sub>2</sub>: Are more advanced Dutch learners of English better able to express sarcastic intonation than intermediate learners of English?*

*RQ<sub>3</sub>: Is there a significant difference between different sentence structures in how effectively Dutch learners of English express English sarcasm?*

The following hypotheses are put forward:

*H<sub>1</sub>: Dutch learners of English will not be able to express accurate sarcastic intonation in English.*

It is expected that Dutch L2-learners of English will not be able to produce effective sarcastic intonation in English, because the intonation of sarcasm is a complex and subtle device (Capelli et al., 1990) and is possibly language-bound (Cheang & Pell, 2013). Furthermore, intonation is not part of foreign language training. Consequently, L2-learners will not have a clear awareness or explicit knowledge on how intonation is used for communicative purposes in English.

*H<sub>2</sub>: More proficient L2-learners of English will be better able to express sarcastic intonation in English.*

The literature review makes a suggestion that L2-acquisition of sarcasm production possibly develops in the same way as L1-acquisition. If this is the case, more advanced learners of English who have had extensive language training can be expected to have had more exposure to sarcastic intonation. This may give them an advantage in sarcasm production.

*H<sub>3</sub>: Sentence structures with less obvious sarcastic cues, namely declarative and exclamatory sentences, will be more difficult to produce in a sarcastic manner for both proficiency groups than tag questions and sentences which include particle discourse markers.*

Tag questions and particle sentences have been identified as sentence structures which can be spoken in a sarcastic manner more easily than others (Kreuz & Gaucci, 2007). Therefore, it is expected that responses of these sentence types will be rated higher on degree of sarcasm for both proficiency groups than declarative and exclamatory sentences. Furthermore, it is expected that nationality of the judges will have no effect on these differences.

## **II. Method**

The data for this study was collected in two experiments. Although only one data set is the result, the experiments are described separately in this section for the sake of clarity.

## Experiment 1

### *Participants*

Fifteen students (11 females, 4 males,  $M_{\text{age}} = 21.3$ ,  $SD = 1.4$ ) of the bachelor programme English Language and Culture and fifteen students (13 females, 2 males,  $M_{\text{age}} = 20.5$ ,  $SD = 1.5$ ) of other programmes at Utrecht University participated in the first stage of this experiment to form two proficiency groups. The intermediate group, referred to as group A, consisted of students which were expected to have an English proficiency level similar to secondary school graduates (B2-level). For this purpose, first, second, and third year students were recruited if they were currently studying a subject which was entirely or almost entirely taught in Dutch. The advanced group, called group B, consisted of third-year students of the English Language and Culture BA course, whose English proficiency level was expected to be near-native (C1-C2-level).

### *Materials*

The stimuli for the first stage of the experiment consisted of thirty pre-recorded sentences describing an event and corresponding verbal responses from an interlocutor. Ten of these sentences were fillers to which the responses were expected to have neutral intonation. The other twenty sentences were used to elicit sarcastic responses. The pre-recorded sentences and corresponding responses were adapted from stimuli and dialogues used in previous studies (Ackerman, 1983; Capelli & Nakagawa & Madden, 1990; Cheang & Pell, 2008; Kreuz & Glucksberg, 1989). There were four different types of responses with possibly different difficulty levels for expressing sarcasm. The four types will be explained below.

Particle sentences included a particle discourse marker at the beginning of the sentence. For example, a particle response would be:

*“Gosh, that'll be fun for a change.”*

The second type was constructed as an exclamation, beginning with the word “what” and ending with an exclamation mark. For example:

*“What a respectful gesture!”*

Declarative sentences had no distinctive features, but included words which have been argued to be cues of sarcasm where possible (Kreuz & Caucci, 2007, 2) in order to decrease the difficulty of these responses. An example of a response is:

*“She sure is a healthy lady.”*

Finally, tag questions were the fourth type of response. For example:

*“You got a great deal, didn't you?”*

Each type consisted of five sentences. The experimental stimuli and fillers are available in Appendix A.

One male native English speaker from the United Kingdom was recruited as the speaker of the pre-recorded sentences. The pre-recorded sentences were meant to aid the participants as much as possible in the production of an effective sarcastic response. Therefore, the speaker was asked to adjust his intonation to convey a sense of mockery or complaint. Since sarcasm has a negative intent, as argued in the literature review, a mocking or whining intonation of the pre-recorded sentences was meant to make it easier for the participants to reply in a natural sarcastic manner.

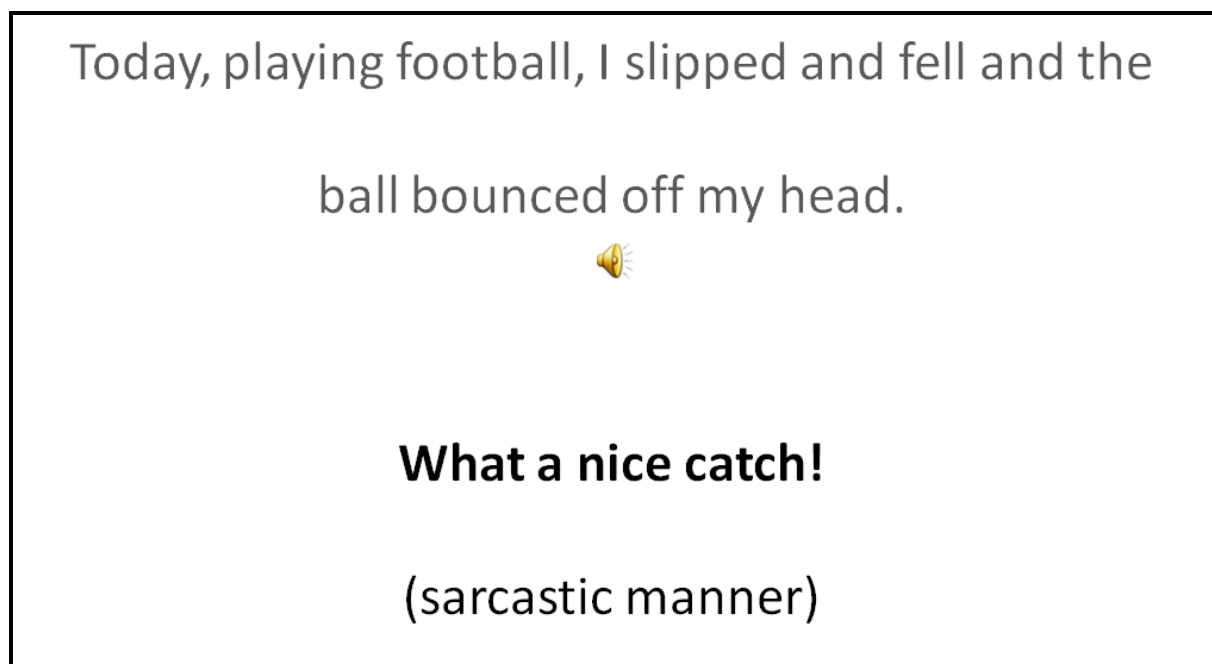
### *Procedure*

The instructions were presented to participants on paper and were written in Dutch (see Appendix B). After reading the instructions, participants were asked to repeat the instructions in their own words. The instructions explained the procedure of the experiment and informed participants that they were allowed to retry the production of responses as often as needed,


until they were satisfied with the manner in which their answer was spoken. They were instructed that in the case of multiple attempts the last response would be used in the experiment, unless participants specifically specified that they preferred a previous attempt. The researcher was present during the experiment, so that when errors were made which changed the meaning of the responses, participants could be asked to repeat a response. Pronunciation errors were not corrected. Participants were also informed that they were allowed to ask questions during the experiment if a sentence or response was unclear. This was done to create a comfortable atmosphere. Beside error corrections and answering questions if necessary, the researcher tried to interfere as little as possible.

There were two orders in which stimuli were presented to participants. These orders were distributed evenly over the participants of the two proficiency groups, so that eight participants of group A and eight participants of group B were presented with the first order, and seven participants of each group were presented with the second order. In both orders the same type of response did not occur twice in a row. Furthermore, each order began with three fillers for practice and ended with two fillers. The remaining neutral fillers were evenly distributed over the twenty sarcastic responses, e.g. a neutral filler was inserted after four sarcastic responses.

Before the experiment began, participants were given the opportunity to practice the procedure with three neutral filler responses. Participants were able to read the orthographic transcription of both the pre-recorded sentence and the response they were supposed to give on a screen, followed by the manner in which the response should be expressed (neutral for the fillers and sarcastic for the stimuli). Image 1 shows the layout of the screen for one of the stimuli.



*Image 1.* Example of the layout

After reading both these sentences, participants would click on the icon which allowed them to listen to the pre-recorded sentence (). Finally, they would reply with the response presented to them. They would then use the right arrow on the keyboard to go to the next stimuli.

Participants were recorded in an audio recording cabin on a Zoom APH-1 H1 audio recording device. Participants were also filmed with a Sony Handycam HDR-CX190 video camera.

## **Experiment 2**

### *Participants*

Participants for Experiment 2 were divided into two groups based on their native language. Three Dutch speakers (2 females, 1 male) of 20, 23, and 26 years old formed a Dutch group. Three British speakers (2 females, 1 male) of 26, 30, and 40 years old formed a British group.

### *Materials*

Experiment 2 was a rating task. Stimuli for this task consisted of the recorded responses of the participants of Experiment 1.

The responses of the participants of Experiment 1 were extracted from the recordings in separate files using *Praat*. If participants had made several attempts at producing a response the last attempt was chosen for Experiment 2, unless otherwise specified by the participants, as they had been instructed. Some recordings contained some background noise, but all responses were intelligible and no recordings were omitted from the experiment.

Neutral fillers were not part of the rating task; only the sarcastic responses were included. The total number of responses to be rated was 600.

### *Procedure*

Rating of the sarcastic responses was done digitally. Participants received an e-mail containing instructions on the procedure of the experiment and on how to download the files needed for the rating task (see Appendix C). The e-mail also contained a link to a Skydrive folder from which all 600 separate audio files and a rating form could be downloaded. The audio files had been renamed so that they were in the right order and no information regarding speaker proficiency could be deduced from the file name. Participants were instructed to play the audio files in chronological order and indicate their rating on the scales on the rating form by highlighting the corresponding score on a 5-point scale (1 = not sarcastic, 5 = very sarcastic). Because of the length of the rating task, participants were advised to take regular breaks. However, they were instructed to finish the rating task in one go, including breaks, rather than spread it out over several days.

Responses were presented to the judges in thirty blocks which corresponded with the thirty participants. The sentences were randomised per speaker in four different orders, which



alternated with each other in varying order. The order of the speakers was also randomised so that speakers of group A and B were intermingled.

### **III. Results**

The data which are being analysed for this study are the rating scores of Experiment 2 for the recordings of Experiment 1. For all statistical tests, an alpha level of .05 was used.

#### *Rater Agreement*

In order to judge the agreement between the different raters, an intraclass correlation coefficient was calculated. The scores of the three Dutch raters were compared with each other and the scores of the British raters were compared with each other as well. The scores were organised per rater. There were 600 items for each rater, because each rater had judged 600 responses.

First, the Dutch raters were compared: rater 1 ( $M = 2.40$ ,  $SD = 1.13$ ), rater 2 ( $M = 2.31$ ,  $SD = 1.18$ ), and rater 3 ( $M = 1.59$ ,  $SD = 0.84$ ). The reliability analysis which was used for testing agreement was a two-way random intraclass correlation coefficient, because both the raters and the speakers whom they had judged had been chosen randomly from a population. The consistency type test was chosen. The test revealed a Cronbach's Alpha of .672, meaning that there was a low to acceptable degree of internal consistency between the raters. The intraclass correlation coefficient for average measures was .672. This is a low coefficient.

Secondly, the British raters were compared using the same test with the same criteria: rater 1 ( $M = 3.19$ ,  $SD = 1.14$ ), rater 2 ( $M = 2.86$ ,  $SD = 1.06$ ), rater 3 ( $M = 3.07$ ,  $SD = 1.31$ ). For the British raters, Cronbach's Alpha was .618. Therefore, internal consistency was acceptable. The intraclass correlation coefficient for average measures was .618.

The conclusion of the inter-rater agreement analysis is that the differences between the raters are low to acceptable for both the British and the Dutch raters. A larger sample of raters might positively influence the agreement between the raters. For this experiment, this was not possible, but future research can take the sample size into account.

### *Rater Language*

Because of the extensive number of rated responses, the mean score of every rater group per sentence type for each speaker was used for analysis. This resulted in eight variables: the mean rating scores for every speaker of the Dutch raters combined and the mean rating scores for every speaker of the British raters combined, for every sentence type.

A mixed design repeated measures analysis of variance was used to investigate significant differences and effects between the different independent variables: proficiency of the speakers, language of the raters, and different sentence types. For the ANOVA, the Skewness and kurtosis statistics indicated that the data for all variables was approximately normally distributed. In addition, Mauchly's test was non-significant, thus indicating that the assumption of sphericity was not violated. Furthermore, Levene's test was non-significant for all dependent variables and Box's test for equal covariance was non-significant as well. Thus, the assumptions of homogeneity and of homogeneity of covariance matrices were assumed.

Table 1 shows the descriptive statistics of the different variables.

Descriptive Statistics

	Proficiency	Mean	Std. Deviation	N
syn1_D	Intermediate	3,0973	,38883	15
	Advanced	2,9860	,41943	15
	Total	3,0417	,40140	30
syn2_D	Intermediate	3,4520	,59559	15
	Advanced	3,7733	,35938	15
	Total	3,6127	,51020	30
syn3_D	Intermediate	2,5560	,58640	15
	Advanced	2,8400	,35124	15
	Total	2,6980	,49641	30
syn4_D	Intermediate	2,6680	,43862	15
	Advanced	2,9567	,43752	15
	Total	2,8123	,45480	30
syn1_B	Intermediate	1,9067	,39171	15
	Advanced	2,1680	,49250	15
	Total	2,0373	,45698	30
syn2_B	Intermediate	2,3807	,50059	15
	Advanced	2,8500	,47422	15
	Total	2,6153	,53526	30
syn3_B	Intermediate	1,7600	,31387	15
	Advanced	2,0927	,29171	15
	Total	1,9263	,34243	30
syn4_B	Intermediate	1,6127	,31329	15
	Advanced	2,0120	,45157	15
	Total	1,8123	,43251	30

Table 1. Descriptive statistics (syn1 = what exclamations, syn2 = particle, syn3 = declarative, syn4 = tag questions, D = Dutch raters, B = British raters)

The first hypothesis states that Dutch learners of English will not be able to express sarcastic intonation in English successfully. If Dutch judges gave higher ratings of sarcasm to the speakers than British judges, the results could be an argument in favour of the hypothesis.

The mixed design repeated measures ANOVA revealed a significant effect for the language of the raters,  $F(1, 28) = 480.600, p < .001, \eta^2 = .95$ . Looking at the means for Dutch and British raters in Table 1, it can be seen that the mean of scores given by British raters are lower than the Dutch mean scores. This means that Dutch raters gave consistently higher sarcasm ratings than the British raters. Pair-wise comparisons revealed that the difference

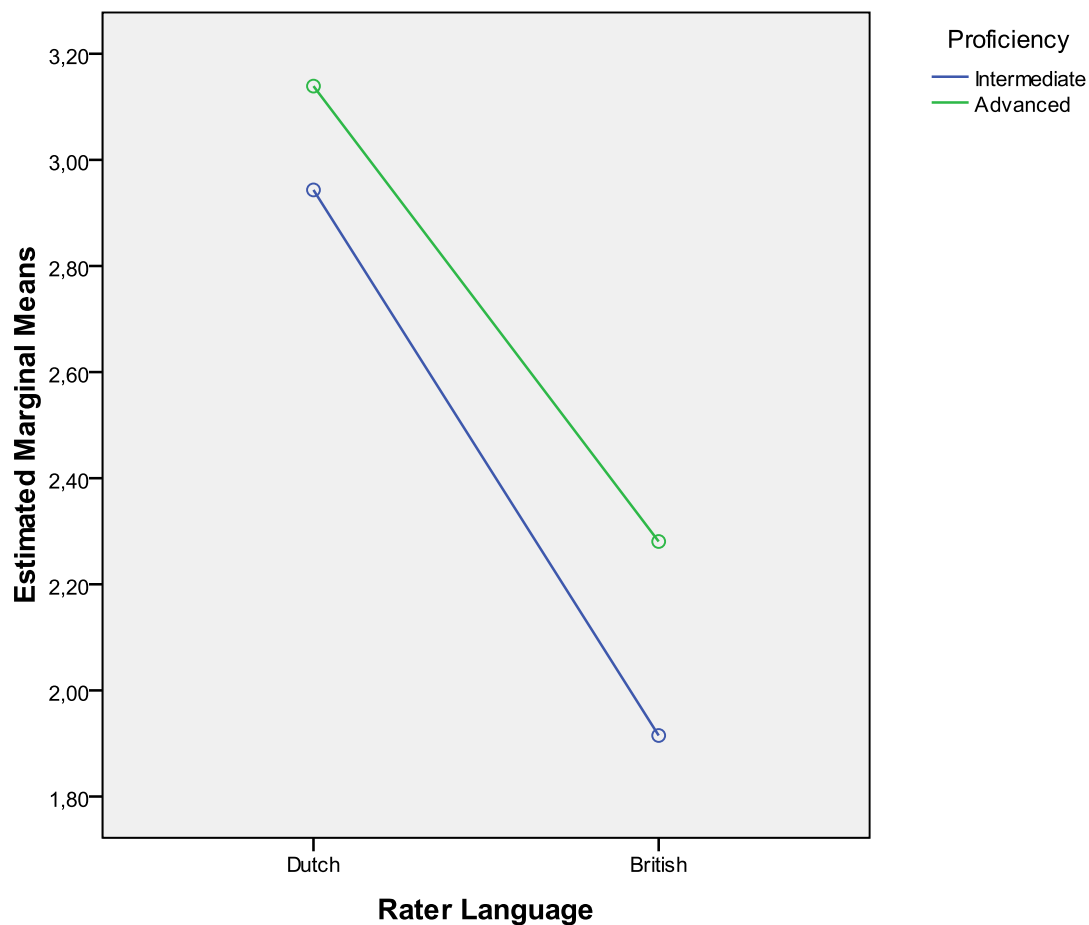
between the means was .943. The outcome of the ANOVA is in accordance with the first hypothesis of this study.

### *Proficiency*

The second hypothesis predicted that speakers with a higher proficiency in English will be better able to express sarcasm in that language. Thus, they would be given higher ratings of sarcasm. The test of between subjects effect of the ANOVA revealed a statistical significance between the two proficiency groups,  $F(1) = 5.553$ ,  $p = .026$ ,  $\eta^2 = .17$ . Pair-wise comparisons and inspection of the mean scores in Table 1 revealed that speakers of the intermediate group were rated lower than speakers of the advanced group ( $MD = .281$ ). Apparently, if scores for both rater groups and all sentence types are considered together, the difference between the scores of proficiency groups is large enough to be significant.

The question remains whether the significant difference between the proficiency groups holds for both the Dutch and the British rater group. The means for the scores of Dutch and British raters for the intermediate speakers were  $M = 2.94$  and  $M = 1.92$  respectively. For the advanced group the means were  $M = 3.14$  and  $M = 2.28$  respectively. The ANOVA showed that there was an effect between rater language and proficiency level of the speakers, but it was only marginally significant,  $F(1, 28) = 3.902$ ,  $p = .058$ ,  $\eta^2 = .122$ . Graph 1 (on the next page) illustrates the distribution of means and shows that the difference between the proficiency groups was larger for the British raters than for the Dutch raters. In order to clarify the outcome of this test, two separate variables were computed with mean scores per speaker per rater group for all sentence types combined. A separate one-way ANOVA was used to assess the difference between the mean scores for intermediate speakers (Dutch raters:  $M = 2.94$ ,  $SD = .45$ ; British raters:  $M = 1.92$ ,  $SD = .27$ ), and for advanced speakers (Dutch raters:  $M = 3.14$ ,  $SD = .29$ ; British raters:  $M = 2.28$ ,  $SD = .35$ ). Levene's test

was non-significant for these variables, thus the equal variances are assumed. In addition, Mauchly's test was non-significant, thus indicating that the assumption of sphericity was not violated. This ANOVA showed a significant difference between the proficiency groups for British raters,  $F(1) = 10.078$ ,  $p = .004$ ,  $d = -1.16$ , but not for Dutch raters,  $F(1) = 2.036$ ,  $p = .165$ ,  $d = -0.52$ . This confirms the statement that the difference between proficiency groups is larger for British raters than for Dutch raters.



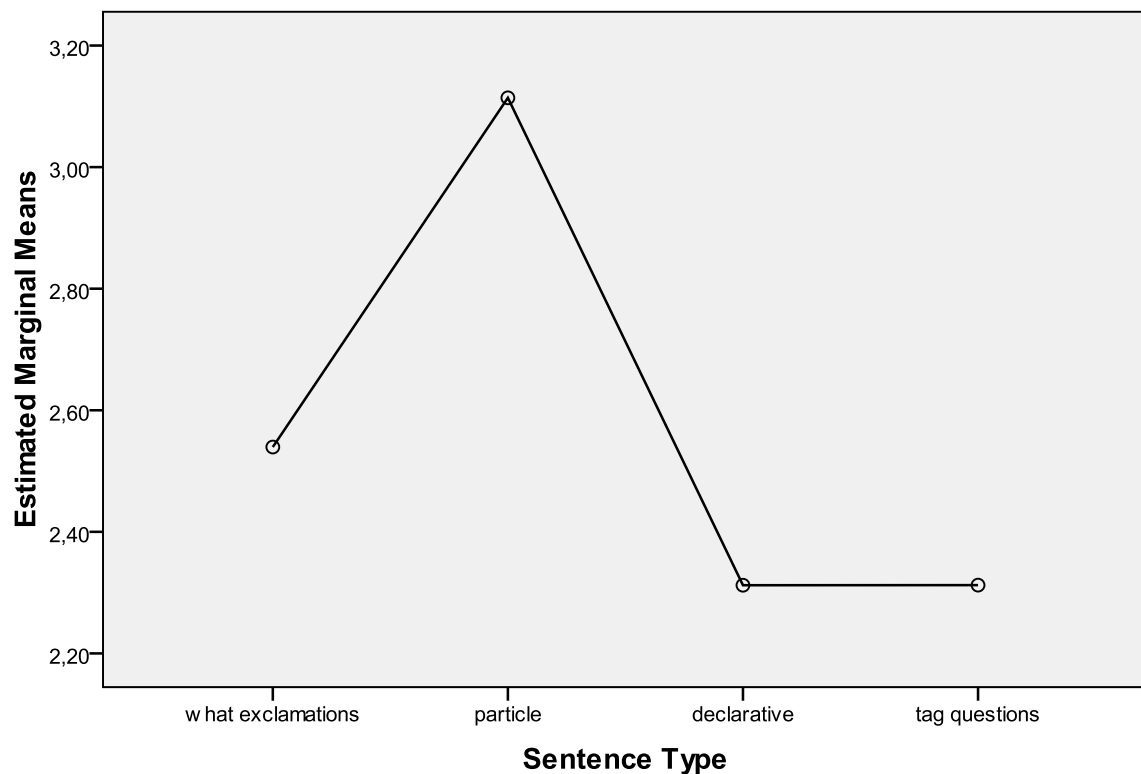
Graph 1. Distribution of means for rater language x proficiency

The second hypothesis states that more proficient speakers would be better able to express sarcastic intonation. This does indeed seem to be the case, as advanced speakers were generally rated higher than intermediate speakers. However, upon closer inspection, the difference between the intermediate and advanced speakers was only significant for the scores

of British raters and not for Dutch raters. It is possible that the ratings of British raters were influenced by other factors. This will be further discussed in the Discussion section below.

### *Sentence Type*

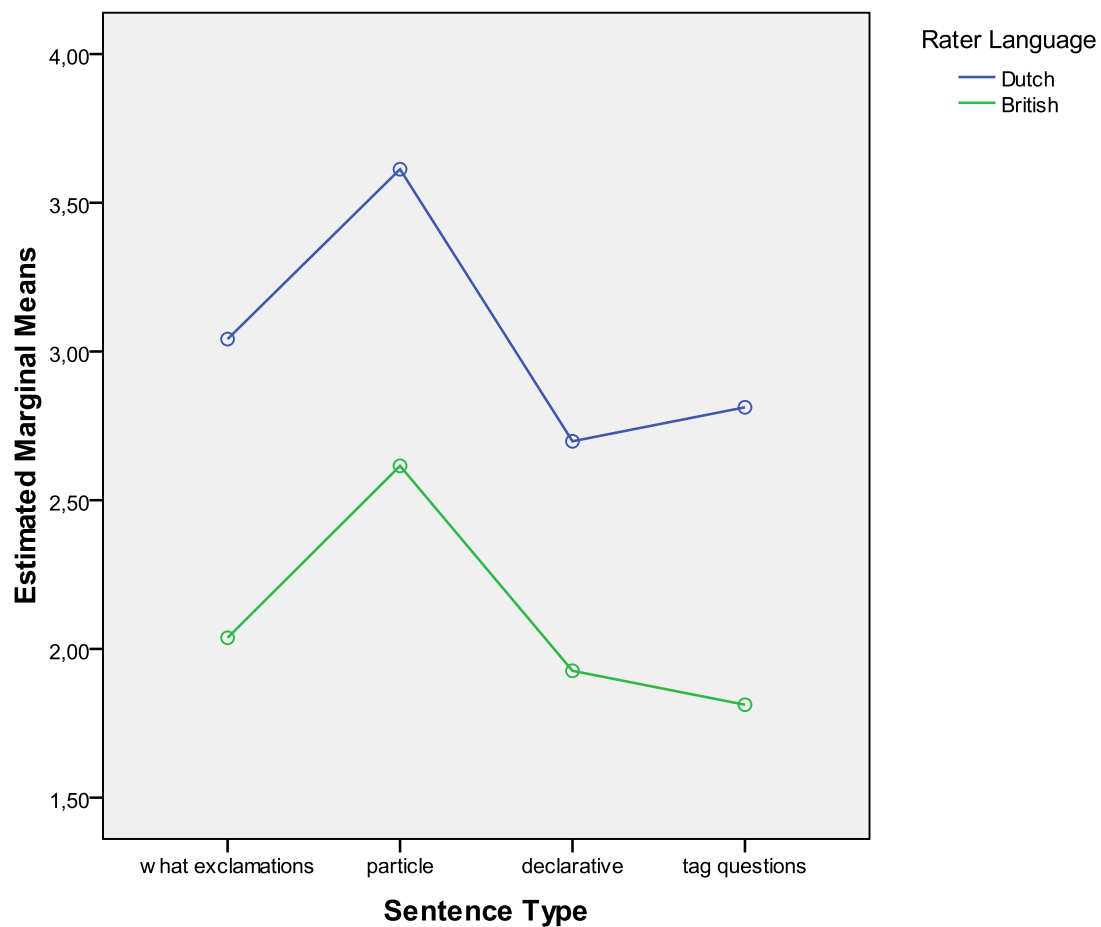
The same mixed design repeated measures ANOVA which was used for analysis above contained results regarding the difficulty of the sentences. There was a significant and large effect for the type of sentence,  $F(3, 26) = 56.066, p < .001, \eta^2 = .87$ . This suggests there is a significant difference between two or more of the sentences. Pair-wise comparisons were used to further analyse this effect. The third hypothesis predicted that declarative and exclamatory sentences would be given lower ratings of sarcasm than particle sentences and tag questions. It became clear from the pair-wise comparisons that exclamatory sentences were rated higher than declarative sentences,  $MD = 0.227, p = .011, 95\% \text{ CI } [0.057, 0.397]$ , and tag questions,  $MD = 0.227, p = .001, 95\% \text{ CI } [0.104, 0.350]$ . However, exclamatory sentences were rated lower than particle sentences,  $MD = -0.575, p < .001, 95\% \text{ CI } [-0.721, -0.428]$ . Particle sentences were rated higher than all other sentences: exclamatory sentences,  $MD = 0.575, p < .001, 95\% \text{ CI } [0.428, 0.721]$ , declarative sentences,  $MD = 0.802, p < .001, 95\% \text{ CI } [0.667, 0.937]$ , and tag questions,  $MD = 0.802, p < .001, 95\% \text{ CI } [0.660, 0.943]$ . Declarative sentences and tag questions were rated lower than the other two sentence types, but there was no significant difference between the two,  $MD = 0.00, p = .998, 95\% \text{ CI } [-0.124, 0.123]$ . Graph 2 (on the next page) provides a visual representation of the described differences between the sentence types for all speakers grouped together.



Graph 2. Distribution of means for sentence type

The test also revealed an effect between language of the raters and sentence type,  $F(3, 26) = 6.841, p = .002, \eta^2 = .44$ . As explained above, Dutch raters gave higher scores of sarcasm to all sentence types than British raters. In addition, the significant difference between the sentence types is maintained in the scores of both Dutch and British raters. This is illustrated in Graph 3 (on the next page).

In addition, the multivariate tests revealed that there was no significant effect between sentence type and proficiency,  $F(3, 26) = 2.466, p = .085, \eta^2 = .221$ . This means that the distribution of mean scores for the different sentence types was roughly the same for both proficiency groups and the difference between the scores for intermediate and advanced speakers was not large enough for significance.



Graph 3. Distribution of means for rater language x sentence type

Finally, mean ratings for every separate particle response were compared to each other. This was done because one extremely high mean rating for a particular sentence of the particle sentence type might have caused the significant difference in the mean score for all particle sentences combined. Shapiro-Wilk, skewness, and kurtosis statistics showed no violation of the assumption of normality. Furthermore, Mauchly's test revealed no violation of the assumption of sphericity. A repeated measures ANOVA was used to compare the five particle responses for both rater groups combined: *"Oh boy; he is a superior chef"* ( $M = 2.80$ ,  $SD = 0.82$ ), *"Gosh, that'll be fun for a change"* ( $M = 3.32$ ,  $SD = 0.63$ ), *"Oh, I guess I better watch out, huh?"* ( $M = 3.39$ ,  $SD = 0.86$ ), *"Gee, aren't you glad you joined the football team"* ( $M = 3.31$ ,  $SD = 0.51$ ), *"Well, it looks like the fish are really biting this year"* ( $M = 2.76$ ,  $SD$



= 0.80). The ANOVA revealed a significant difference between the responses,  $F(4, 20) = 5.53, p = .004, \eta^2 = .53$ . Pair wise comparisons indicated that two sentences were rated significantly lower than the others: “*Oh boy; he is a superior chef*” and “*Well, it looks like the fish are really biting this year*”.

## VI. Discussion

The present study aimed to investigate the ability of second language learners of English to produce sarcastic intonation in English and the influence of different sentence structures on this ability. Scores for the degree of sarcasm for every speaker of by Dutch and British raters formed the data by which the hypotheses were tested.

### *Ability*

It was hypothesised that the speakers in this experiment would receive higher ratings of sarcasm from Dutch judges than British judges. Because Cheang and Pell (2013) argue that intonation of sarcasm cannot be perceived in another language, it is possible that it cannot be produced in another language either, in which case the Dutch judges would recognise the sarcastic intonation of the Dutch speakers, as it is familiar to them, and the British judges would not. The current study found that this seemed to be the case. Dutch judges gave significantly higher ratings of sarcasm than British judges. The mean ratings for speakers by Dutch raters ranged from  $M = 2.03$  to  $M = 3.50$ , whereas the mean ratings of British raters ranged from  $M = 1.45$  to  $M = 2.93$ . These findings corroborate the ideas of Cheang and Pell (2013), who suggested that every language has a different intonation pattern for sarcasm, thereby rejecting the idea of universal prosodic cues for sarcasm, although such an idea had been suggested by a number of authors (Attardo & Eisterhold & Hay & Poggi, 2003; Creusere, 1999; Nakassis and Snedeker, 2002; Rockwell, 2000, 2007). In this way, sarcasm

seems to differ from other emotions which have been argued to be recognisable across languages. Morton and Trehub (2001) name happiness and sadness as perceivable in foreign languages, even by children. Perhaps the difference between sarcasm and these emotions lies in the suggestion that sarcasm is a verbal mechanism to convey other emotions and intentions such as humour, irony, anger, or contempt (Kreuz & Glucksberg, 1989), rather than an emotion itself. If the intonation of sarcasm in English indeed differs from the Dutch intonation, the results of this study suggest Dutch L2-learners of English lack proficiency in English prosody. Consequently, the responsibility lies with educational research to investigate the effectiveness of prosodic language training in this area.

### *Proficiency*

The second question in this study is whether there is a difference between less proficient and more proficient L2-speakers in how effectively they are able to produce sarcasm in English. It was expected that more proficient speakers would be given higher ratings of sarcasm. If this was the case, it can be argued that L2-acquisition of sarcasm production develops in a similar way as L1-acquisition, namely that it improves with age, exposure to, and proficiency of the language. However, if sarcasm production does not develop in this way and needs to be trained in order to be developed, it must be considered that sarcastic intonation training does not yet exist. Consequently, even advanced learners of English would not be able to produce the intonation of sarcasm effectively in their second language.

The results of this study do not provide an entirely clear answer to the possibilities which have been described. A significant difference between proficiency groups was discovered when testing the ratings of all raters combined. However, the difference did not occur when the different types of sentences were tested. In addition, there was a significant difference between the ratings for intermediate and advanced speakers of British raters, but

this was not the case for the Dutch raters. As the sample sizes were smaller in these divided tests, a loss of significance may have occurred. However, there may be other explanations. The mean difference between Dutch and British ratings for the intermediate group was  $MD = 1.02$ . For the advanced group this was  $MD = 0.86$ . Thus, it seems the British raters were harder on the intermediate speakers than on the advanced speakers in comparison with the Dutch raters. Possible explanations are that the intermediate speakers had a more distinct foreign accent in their speech. British may have let this accent influence their judgment on how sarcastic the responses sounded, whereas Dutch raters can be considered to be more used to a Dutch accent in English. Therefore, this may not have been an influencing factor for them. Furthermore, the intermediate speakers seemed to have greater trouble producing the responses because their general proficiency was lower than the more proficient speakers, who were more comfortable in pronouncing the English sentences. Hence, a possible explanation for a difference between the proficiency groups is that the task was more difficult for speakers of group A, because of their general proficiency level. However, this does not explain why the difference should only be significant for British rating scores. Unfortunately, the outcome of the tests and the factors which possibly influenced the results means that in this study no definite answer can be given on the question whether the ability to express sarcastic intonation in English increases when L2-proficiency does.

However, it is somewhat surprising that the advanced speaker group, considered to be at a near-native proficiency level of English, did not receive higher marks from the British judges, but were rated higher by Dutch judges, as mentioned in the previous section. If the hypothesis was true, sarcasm production should improve as learners are more exposed and trained in the English language. Considering this, the advanced group should have attained a near-native sarcastic intonation in English because of the advanced language training they have received. As a result, there should have been no difference between Dutch and British

ratings for the advanced group, yet this does not seem to be the case. The results suggest that, although the advanced group is possibly better at producing sarcasm than the less proficient group, they still retain a partly Dutch intonation of sarcasm, prompting the Dutch judges to give higher ratings than the British judges. However, it should be taken into account that no control group of British speakers provided responses for the rating task. Therefore, the advanced group cannot be compared to native speakers of English, thus no conclusions can be drawn on whether or not the group with advanced proficiency in English has a native-like proficiency in sarcasm intonation.

### *Sentence Types*

Finally, the difficulty of different types of sentences was investigated. Considering the available literature on lexical influences in sarcasm (Kreuz & Gaucci, 2007), it was expected that of the four types of sentences in the experiment, responses with particle discourse markers and tag questions would be easier to produce than declarative and exclamatory sentences. The current study found that particle responses were given significantly higher ratings for both judging groups than the other three types of responses. From the other three sentence types, exclamatory sentences were rated significantly higher than declarative sentences and tag questions, between which no significant difference was found.

Dutch judges rated all sentences higher than British judges, continuing the trend which has been described above, but the distribution between the different types of sentences is relatively very similar, with particle responses receiving the highest ratings for both groups.

Although it is surprising that tag questions were not rated higher than declarative and exclamatory sentences, the lower difficulty of particle responses was predicted. According to Kreuz and Gaucci (2007), particle discourse markers can carry such sarcastic significance that the words themselves may be considered as cues for a sarcastic interpretation (2). The

sentences “*Oh boy; he is a superior chef*” and “*Well, it looks like the fish are really biting this year*” were rated lower than the other three responses in this type group: “*Gosh, that'll be fun for a change*”, “*Oh, I guess I better watch out, huh?*”, “*Gee, aren't you glad you joined the football team*”. Kreuz and Gaucchi (2007) identified “gee” and “gosh” as particle discourse markers which could serve as cues for sarcasm, but the other particles used in the sentences of this type, “oh boy,” “well,” and “huh” were not discussed in their study. Hence, whether they are clear cues of sarcasm as well is unclear. A possible explanation for the differences between these sentences of the same type is that some particle discourse markers may not carry the same sarcastic implications as “gee” and “gosh”.

A possible explanation for the ratings of exclamatory responses, which were higher than declarative sentences and tag questions, is not given in the literature we have reviewed. A speculative explanation is that the structure “*What ...!*” is commonly used as a sarcastic construction in contemporary film and theatre. However, there is no empirical data to support that statement.

### *Limitations*

In order to accurately interpret the results of the current research, a number of limitations need to be considered. Firstly, because of time limitations, the number of participants who acted as raters in the experiment was relatively small. Because the intonation of sarcasm may differ per country, region, and possibly even per individual, the results would have been more representative if a larger number of judges had been used. The experimental method and responses of Experiment 1 could well be used for data gathering among a larger sample of raters. Secondly, the experiment contained no control group of British speakers for a native representation of the responses. A control group of British speakers would have provided the opportunity to compare the advanced speaker group with a native sample. Thirdly, the

sarcastic responses which were recorded in this experiment were not spontaneous, although attempts were made for the production to be as spontaneous as possible. Rockwell (2000) suggests that there is a difference between posed and spontaneous sarcasm, arguing that posed sarcasm is easier to perceive. Because of this, raters may have given higher scores of sarcasm than they would have given if the speech had been spontaneous. As the conditions were the same for all speakers in the experiment, it is of no direct influence on the results of our study. However, the distinction between spontaneous and posed sarcasm should be taken into account making generalised statements about the outcome of the research. Finally, the difference between production and perception of sarcasm, which the literature has brought to attention, should be considered in this experiment. Participants of Experiment 1 produced the responses, but whether or not their sarcastic intonation was perceived by the raters of Experiment 2 depends not only on the participants' ability to produce, but also on the judges' ability to perceive sarcasm by use of intonation. For more definite answers, the size of the rating groups should be increased, so that a variation between perception abilities of participants might be minimised.

## **V. Conclusion**

The present study was designed to determine whether Dutch L2-learners of English are able to express sarcastic intonation in English. Furthermore, the aim was to assess possible lexical factors which may be of influence on how effectively they are able to do so. The findings suggest that in general Dutch learners of English are not able to produce native-like intonation of sarcasm in English. However, particle discourse markers in a sentence made it easier to produce sarcasm. Exclamatory sentences were also somewhat easier to express in a sarcastic intonation. Returning to the hypotheses posed at the beginning of this study, it is now possible to state that, although much is still unclear on the topic, there are strong arguments that

intonation of sarcasm differs per language and cannot be learned by exposure to a foreign language alone.

The current findings add to a small but increasing body of literature on cross-lingual intonation studies. Further research needs to be undertaken before an association between proficiency level in the foreign language and the ability of producing sarcastic intonation can be established. Several questions remain unanswered at present. In future investigations it might be interesting to use larger sample sizes and spontaneous sarcastic remarks drawn from conversations instead of scripted sarcastic responses, as it would enhance the generalisability of the arguments made in this study.

### **Acknowledgments**

This research was conducted as a bachelor's thesis. I would like to thank Aoju Chen, my thesis supervisor, for her feedback and tips which made this thesis into a learning process that I look back on with pride. I would also like to thank everyone who participated and helped me recruit other participants in the Netherlands and the UK.

## References

- Ackerman, B.P. Form and Function in Children's Understanding of Ironic Utterances. (1983). *Journal of Experimental Child Psychology*, 35, 487-508. Retrieved from <http://www.sciencedirect.com/>
- Attardo, S., & Eisterhold, J., & Hay, J., & Poggi, I. (2003). Multimodal Markers of Irony and Sarcasm. *Humor*, 16(2), 243-260. Retrieved from EBSCOhost Academic Search Elite.
- Bowes, A., & Katz, A. (2011). When Sarcasm Stings. *Discourse Processes*, 48, 215-236. doi:10.1080/0163853X.2010.532757
- Capelli, C.A., & Nakagawa, N., & Madden, C.M. (1990). How Children Understand Sarcasm: The Role Of Context and Intonation. *Child Development*, 61, 1824-1841. Retrieved from EBSCOhost Academic Search Elite.
- Cheang, H.S., & Pell, M.D. (2008). The Sound of Sarcasm. *Speech Communication*, 50, 366-381. doi:10.1016/j.specom.2007.11.003
- Cheang, H.S., & Pell, M.D. (2013). Recognizing Sarcasm Without Language: A Cross-Linguistic Study of English and Cantonese. In S. Attardo & M.M. Wagner & E. Urios-Aparisi (Eds.), *Prosody and Humor* (pp. 15-36). Amsterdam & Philadelphia: John Benjamins B.V.
- 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/>
- Creusere, M.A. (2000). A Developmental Test of Theoretical Perspectives on the Understanding of Verbal Irony: Children's Recognition of Allusion and Pragmatic Insincerity. *Metaphor and Symbol*, 15(1-2), 29-45. doi:10.1080/10926488.2000.9678863



- Gibbs, R.W., & O'Brien, J. (1991). Psychological Aspects of Irony Understanding. *Journal of Pragmatics*, 16, 523-530. Retrieved from <http://www.sciencedirect.com/>
- Katz, A.N., & Blasko, D.G., & Kazmerski, V.A. (2004). Saying What You Don't Mean: Social Influences on Sarcastic Language Processing. *Current Directions in Psychological Science*, 13(5), 186-189. doi:10.1111/j.0963-7214.2004.00304.x
- Kreuz, R.J. (2000). The Production and Processing of Verbal Irony. *Metaphor and Symbol*, 15(1-2), 99-107. doi:10.1080/10926488.2000.9678867
- 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 <http://www.upenn.edu/>
- 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. & Bert-Erboul, A. (2005) French-Speaking Children's Understanding of Sarcasm: The Role of Intonation and Context. *Journal of Speech, Language, and Hearing Research*, 48, 610-620. Retrieved from EBSCOhost Academic Search Elite.
- Lee, C.J., & Katz, A.N. (1998). The Differential Role of Ridicule in Sarcasm and Irony. *Metaphor and Symbol*, 13(1), 1-15. doi:10.1207/s15327868ms1301\_1
- Morton, J.B., & Trehub, S.E. (2001). Children's Understanding of Emotion in Speech. *Child Development*, 72(3), 834-843. Retrieved from <http://www.utoronto.ca/>
- Muecke, D.C. (1980). *The Compass of Irony*. London: Methuen.
- Nakassis C., & Snedeker C. (2002) Beyond sarcasm: Intonation and Context as Relational Cues in Children's Recognition of Irony. In A. Greenhill & M. Hughs & H. Littlefield & H. Walsh (Eds.), *Proceedings of the Twenty-sixth Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

- Recchia, H.E., & Howe, N., & Ross, H.S., & Alexander, S. (2010). Children's understanding and production of verbal irony in family conversations. *British Journal of Developmental Psychology*, 28, 255-274. doi:10.1348/026151008X401903
- Rockwell, P. (2000). Lower, Slower, Louder; Vocal Cues of Sarcasm. *Journal of Psycholinguistic Research*, 29(5), 483-495. Retrieved from <http://www.springer.com/>
- Rockwell, P. (2001). Facial Expression and Sarcasm. *Perceptual and Motor Skills*, 93, 47-50. Retrieved from <http://www.amsciepub.com/>
- Rockwell, P. (2007). Vocal Features of Conversational Sarcasm: A Comparison of Methods. *Journal of Psycholinguistic Research*, 36, 361-369. doi: 0.1007/s10936-006-9049-0

## Appendices

### A. Stimuli and Fillers

<i>Sentence type</i>	<i>Pre-recorded remark</i>	<i>Participant's response</i>
Particle	My father had made a five course Christmas dinner; the next day we were all sick with food poisoning.	Oh boy; he is a superior chef.
Particle	Tomorrow's class is going to be about plants again.	Gosh, that'll be fun for a change.
Particle	I heard Peter, that skinny kid with glasses, is going to beat you up after school.	Oh, I guess I better watch out, huh?
Particle	My first football training coach made us run 5 miles; some of the guys threw up.	Gee, aren't you glad you joined the football team.
Particle	I went fishing but didn't catch anything.	Well, it looks like the fish are really biting this year.
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 a spectacular result!
What exclamation	Today, playing football, I slipped and fell and the ball bounced off my head.	What a nice catch!
What exclamation	My piano performance has been cancelled.	What a shame!
What exclamation	My sister's piano performance went very poorly.	What a surprise!
Declarative	That horrid woman smokes a pack a day.	She sure is a healthy lady.
Declarative	I bought a new game and I thought it would be too hard, but I learned it in five minutes.	It takes a real genius to learn that game.
Declarative	I went for a run and I came back dripping wet.	It's a wonderful day out.
Declarative	I bought you another pair of socks for Christmas.	Socks are just what I need.
Declarative	My plane was an hour late.	Those punctual airlines are really impressive.
Tag question	I put my homework off for two hours, but then it only took ten minutes.	That sure took a lot of effort, didn't it?
Tag question	I think I didn't even get one right on that test.	You really did well this time, didn't you?
Tag question	My little sister kicked me in the shins.	Your sister is very sweet, isn't she?
Tag question	I traded my cricket bat for a toy truck, but now I find out it's broken.	You got a great deal, didn't you?
Tag question	My brother wanted to help me move and he dropped my grandfather's clock.	He's really helping you out, isn't he?
Filler	My favourite colour is green.	Mine is purple.
Filler	My brother was accepted to the police academy.	I know, your sister told me.
Filler	I'm going for a run.	Good luck!
Filler	I don't know much about politics.	Neither do I.
Filler	Thank you for the flowers.	You're welcome.
Filler	You shouldn't buy things on those dodgy websites.	I agree.
Filler	Shall I call you tomorrow and let you know the time and place?	Yes, please.
Filler	My mother makes the best apple pie.	She's a very good baker.
Filler	Do you know if he will be there on Saturday?	I don't think so.
Filler	Avoiding the main roads will save you a lot of time.	That's right.

### *B. Instructions Experiment 1*

Leuk dat je mee wilt doen aan dit experiment!

#### **Wat is jouw taak?**

Stel je voor: je wordt opgebeld door een goede vriend. Je kletst een tijdje met elkaar. Dan maakt je vriend een opmerking.

Het is de bedoeling dat je antwoord geeft op je vriend met de tekst die op de dia staat. De manier waarop je moet antwoorden is ook gegeven. Het belangrijkste is dat jouw antwoord zo natuurlijk mogelijk klinkt. Er is dus geen goed of fout.


#### **Hoe werkt het experiment?**

Op het scherm komen zo meteen drie zinnen te staan:

Opmerking van je vriend (in grijs)

**Het voorgeschreven antwoord dat jij gaat geven** (dikgedrukt)

(De manier waarop jij antwoord geeft) (tussen haakjes)

Je krijgt eerst de tijd om de opmerking en jouw antwoord helemaal te lezen. Daarna kun je de opmerking van de vriend die jou belt beluisteren door op het icoontje  te klikken. Dit mag je zo vaak doen als je wilt.

Vervolgens spreek je jouw antwoord op de aangegeven manier uit. Ook dit mag je zo vaak proberen als je wilt.

Als je tevreden bent over jouw antwoord, klik je op het pijltje naar rechts → op het toetsenbord om naar de volgende zin te gaan.

#### **Waar moet je nog meer op letten?**

Let erop dat je de antwoorden precies zegt zoals ze op het scherm staan. Dus probeer er geen woorden zelf bij te verzinnen.

#### **Ten slotte,**

Als je vragen hebt, aarzel dan niet om ze te stellen, ook tijdens het experiment.

Veel plezier en succes!

### C. Instructions Experiment 2

Dear raters,

First of all, thank you for agreeing to help with this rating experiment!

#### What is your task?

You will be presented with a number of very short audio files. The audio files are responses of a group of speakers to remarks of their friend. The speakers were supposed to respond in a sarcastic tone. Your task is to rate how sarcastic each response sounds to you on a 5-point scale. The score “1” stands for “not sarcastic”. The score “5” stands for “very sarcastic”.

#### How does it work?

First, please click on this link: <https://skydrive.live.com/redirect?resid=6FFD05174009A07E%21107> download the audio and two Word files needed for the rating onto your computer (select “download” and “Download as .zip”).

The audio files can be put into a playlist in a media player of your choice, so that you do not have to open each file separately. The name of each audio file corresponds to the speaker code and response number. For example: the audio file “Speaker 01 response 01.wav” corresponds to response 01 of speaker 01.

One Word file is the questionnaire on your language and education background for you to fill in.

The other Word file “Rating Task Form.doc” is where you specify your ratings. The form consists of 30 blocks. Each block contains the responses from one speaker together with the corresponding remarks. There are five trials per page. An example of a trial is as follows:

01.

Remark: *My mother-in-law always smirks and snorts loudly when I misspeak.*

Response: *What a respectful gesture!*

not sarcastic

1

2

3

4

5

very sarcastic

This is the orthographic transcription of the response. Please note that it can happen that the speaker occasionally mispronounced or added a word. This, however, should not influence your rating of the sarcastic tone.

Please fill in the rating form electronically by opening the rating form in Word and highlighting the corresponding score in the scale, as shown in the example.

Please note that it is important that you finish rating the sentences from one speaker before moving on to another speaker. There are 30 speakers in total.

Please listen to each response no more than three times before you give a score.

Finally, it may take an hour to 90 minutes to rate all the sentences, including short breaks. Please feel free to take short breaks during the rating experiment. Please complete the rating in one go (including short breaks), instead of spreading the work over several hours or days.

Please send the filled rating form and questionnaire to [D.M.J.dejong@students.uu.nl](mailto:D.M.J.dejong@students.uu.nl) by the 18<sup>th</sup> of January. If this is not feasible for you, please let me know.

Good luck, have fun, and thank you very much!