



Prosodic cues of sarcastic speech in French: slower, higher, wider

Hélène Lævenbruck^{1,2}, Mohamed Ameer Ben Jannet^{1,3,4,5}, Mariapaola D'Imperio^{6,7},
Mathilde Spini⁶, Maud Champagne-Lavau⁶

¹GIPSA-lab, Speech & Cognition Department, UMR CNRS 5216, Univ. Grenoble-Alpes, ²LPNC, UMR CNRS 5105, Univ. Grenoble-Alpes, ³Lne, Paris, ⁴LIMSI UPR 3251, Paris, ⁵LPP UMR 7018, Paris, ⁶Aix-Marseille Université, CNRS, LPL, UMR 7309, Aix-en-Provence, France, ⁷Institut Universitaire de France

Helene.Loevenbruck@gipsa-lab.grenoble-inp.fr, mariapaola.dimperio@lpl-aix.fr,
maud.champagne-lavau@univ-amu.fr

Abstract

Verbal irony is characterized by the use of specific acoustic modulations, especially global prosodic cues as well as vowel hyperarticulation. Little is known concerning the expression of sarcastic speech in French. Here we report on global prosodic features of sarcastic speech in a corpus of declarative French utterances. Our data show that sarcastic productions are characterized by utterance lengthening, by increased f_0 modulations and a global raising of the pitch level and range. The results are discussed in the light of results on the acoustic features of ironic speech in languages other than French.

Index Terms: sarcastic speech, French, intonation, prosody.

1. Introduction

Irony and sarcasm are the most prevalent forms of non-literal communication in our culture. Verbal irony is a mode of expression in which what is stated differs from (or is even opposed to) what is meant. Irony exists in the majority of the languages and cultures of the world [22]. For an ironic meaning to be conveyed, the communicative effect is based on success on the part of the listener to understand the ironic intent of the utterance, i.e. the incongruity between the literal and the intended meaning [16]. Irony can convey a positive or negative meaning. Sarcasm is considered as a subtype of irony, which conveys a negative, critical or hostile meaning.

Irony cues can in fact be conveyed lexically (e.g., use of exaggerated adjectives and adverbs), nonverbally (e.g., facial expressions) and through prosodic modulations [4, 6; *inter alia*]. Previous research has shown that in different languages ironic speech is acoustically differentiated from literal speech, and these modulations are assumed to aid the listener in the comprehension process by acting as cues that mark utterances as ironic [6, 8]. In particular, several research sources have highlighted the importance of prosody (intonation, rate/rhythm, phrasing) as a cue for detecting sarcasm [10].

Some researchers have also proposed that acoustic irony cues are only employed if the common ground is not sufficient to indicate the intended message [11]. Note, though, that more recent research has shown that ironic content can be identified even in absence of contextual cues thanks to global acoustic/prosodic cues [6]. It has also been shown that young children can recognize the intonational markers of sarcasm, and this ability is developmentally distinct from the ability to recognize sarcasm through semantic or contextual cues [1, 2; *inter alia*]. However, we still do not know what is the actual role of prosody, in particular of intonational phonology features [17], in irony comprehension.

Another line of research claims, on the contrary, that irony is not associated with a particular intonational contour and that it is thanks to a multitude of cues other than intonation, including extralinguistic information, that listeners manage to recognize that a statement is ironic [7]. For instance, it has been shown that several factors, such as the degree of incongruity between context and speaker utterance can influence the extent to which ironic intent is perceived [6, 16].

Concerning, the actual acoustic cues, sarcasm appears to be encoded in speech through various global manipulations in acoustic parameters such as fundamental frequency (f_0), amplitude, speech rate, voice quality and vowel hyperarticulation [4, 7, 11, 12, 15, 25; *inter alia*]. However, owing to methodological differences across studies, the available data are quite controversial, and the relative importance of particular acoustic parameters for signalling sarcasm and their directionality cannot be fully determined.

An additional problem comes from the fact that gradual prosodic variability is modulated through a phonological structure of intonation that is language-specific. Finally, local and global duration manipulations are usually conflated in existing data, rendering the results difficult to evaluate. In this study, we explore the expression of sarcasm in French, for which data are still lacking. We will specifically examine the hypothesis that sarcastic utterances which were correctly identified as being sarcastic are globally lengthened and that their f_0 level is either lowered or raised as what was found by some studies in a number of Germanic languages.

2. Method

2.1. Production task

2.1.1. Material

48 utterances distributed in two attitude conditions (24 sarcastic utterances, 24 literal utterances) were used. To induce sarcastic or literal attitude, each sentence followed a short context as in the example (Table 1). Short stories were adapted from [9]. This method allowed us to place speakers in pragmatic situations naturally inducing sincere or sarcastic utterances. Thus, depending on the preceding context, the last sentence was produced as being marked either by a sarcastic or by a sincere tone. All contexts were recorded by a female native speaker of French. Audio signals were recorded using a digital recorder, in a soundproof room and were played through loudspeakers to the participants.

Table 1. Example of contexts used to elicit literal and sarcastic utterances

	Sarcastic condition	Literal condition
Context	<p>Emilie voit Pierre arriver au travail le lundi matin. Il est pâle et a l'air d'avoir très mal dormi.</p> <p><i>'Emilie sees Peter arriving at work on Monday morning. He is pale and seems to have not slept well.'</i></p>	<p>Emilie voit Pierre arriver au travail le lundi matin. Il est resplendissant et prêt à commencer la semaine d'un bon pied.</p> <p><i>'Emilie sees Peter arriving at work on Monday morning. He is radiant and ready for a great new week.'</i></p>
Target sentence	<p>Il est en pleine forme</p> <p><i>'He is in great shape.'</i></p>	<p>Il est en pleine forme</p> <p><i>'He is in great shape.'</i></p>

2.1.2. Participants and Procedure

Twelve native speakers of French (6 females, 6 males) were recruited for the production task. They were between 23 and 36 years old (mean: 29± 5.2). They were graduate students or faculty members of GIPSA-lab. None of them had any known speech or hearing problems and they were naïve with respect to the purpose of the experiment.

Speakers were instructed to listen to each of the recorded contexts and then to read the target sentence on a computer screen so that it would fit the preceding context (either with a sarcastic or a neutral tone of voice). Utterances were recorded with an AKG microphone, the audio signal was sampled at 44100Hz. Stimuli were distributed into two blocks, the first one containing the stories inducing sarcasm, the second one containing the stories inducing a literal reading.

2.2. Stimulus validation

A perception task was performed to ensure that stimuli produced with each targeted tone (sarcastic, sincere) were prototypical of each category.

2.2.1. Material

234 pairs (sarcastic, sincere) of recorded utterances from the initial 288 pairs (12 speakers x 24 sentences) obtained in the production task were used for the stimulus validation task. The productions of 2 speakers (1 female, 1 male) had to be discarded due to noise as well as a few other utterances containing pronunciation errors. Before the validation test, all stimuli were normalized in mean intensity using a *Matlab* script since unequal loudness across speakers might have influenced the test.

2.2.2. Participants and procedure

Twenty native speakers of French (10 females, 10 males) were recruited at GIPSA-lab for the perception task. They were between 21 and 62 years old (mean: 30± 9.2) and were recruited from the same population as the speakers. None of them had any known speech or hearing problems and they were naïve with respect to the purpose of the experiment.

Participants listened to utterances produced in the production task without any previous context. They were asked to judge if

the target sentence was sarcastic or sincere in a two-alternative forced choice procedure. They also had to rate their confidence level in interpretation using a 5-point Likert scale. Given the amount of stimuli presented, the data were divided in two blocks, each one containing utterances produced by 5 speakers. Each block (of 228 and 240 sentences respectively) was then evaluated by 10 participants.

Overall the accuracy score (the percentage of correct answers for all of the 468 utterances for all participants) was quite high (79%). This score is higher than the 67% found by [6] in a similar procedure for English or that obtained by [20] for French (around 70%). Hence our results suggest that sarcastic and literal utterances can be distinguished even without context.

To ensure that acoustic analyses were carried out on prototypical utterances that were robustly identified by a range of participants, pairs of utterances were kept for further acoustic analyses only if each of the utterances in the pair (literal and sarcastic) had been correctly identified. Utterances were considered as correctly identified when 70% of the listeners had categorized them correctly, with a confidence level of 4 or above. Out of the original 234 pairs, a total of 104 utterances (52 pairs) from 9 speakers were validated, 64 of which had been pronounced by male speakers (40 by female speakers).

2.3. Acoustic analysis

All validated utterances were acoustically analyzed using *Praat* [5]. Pitch level was computed as the mean of the f_0 values transformed in semitones (relative to 100 Hz) for the utterance as a whole, in order to examine whether level is a cue to sarcasm. Pitch span (in semitones) was computed by subtracting the minimum from the maximum f_0 value (in semitones) for each utterance as a further index of pitch variation. Duration (in s) was computed for each utterance to examine whether sarcasms is associated with any rhythmic effect. Percent lengthening from the literal to the sarcastic version was computed by subtracting the duration of the literal version from that of the sarcastic version and then dividing this difference by the duration of the literal version.

3. Results

3.1. Overall results

Figure 1 provides the superposition of literal and sarcastic f_0 contours for each validated utterance. Note that, from a first informal inspection of the contours, sarcastic f_0 contours seem to show a higher pitch level and a wider span on average. Sarcastic utterances also seem to be longer. The aim of the acoustic analyses presented below was to verify our preliminary observations through quantitative measurements.

An example of neutral and sarcastic versions of one sentence is given in Figure 2, in which f_0 is on average higher in the sarcastic version. Also, a slight final rise (M%) is observed in the sarcastic version, instead of the falling contour (L%) observed in the literal version. The sarcastic version was also longer.

3.2. Pitch level and span

Pitch level and span (in semitones) were computed for each utterance. Table 2 provides means and standard deviations (in

parentheses) for level and span for literal and sarcastic utterances, in all validated pairs. Two paired-samples t-test were conducted on all utterances to compare f_0 level and f_0 span in literal and sarcastic conditions.

There was a significant difference in mean f_0 for literal ($M=5.93$, $SD=6.05$) and sarcastic ($M=6.86$, $SD=5.65$) conditions; $t(51)=-2.92$, $p = 0.005$. Also, there was a significant difference in span for literal ($M= 10.89$, $SD=4.31$) vs. sarcastic ($M=13.29$, $SD=4.55$) conditions; $t(51)=-2.47$, $p = 0.01$. These results suggest that sarcasm has an effect on pitch level and span. Specifically, as illustrated by Figure3, our results suggest that when producing a sarcastic utterance, pitch level is higher and the pitch span is expanded.

Table 2. Mean and standard deviation (in parentheses) of f_0 level and span in literal and sarcastic utterances

	f_0 level (st)	f_0 span (st)
Literal	5.93 (6.05)	10.89 (4.31)
Sarcastic	6.86* (5.65)	13.29* (4.55)

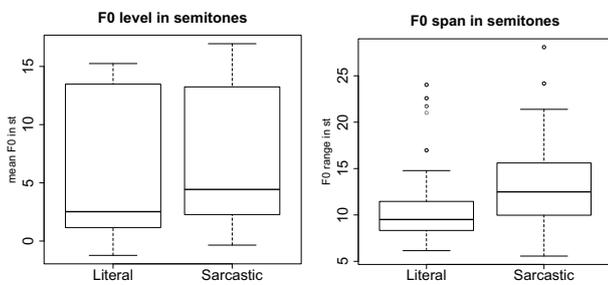


Figure 3: Boxplots of the f_0 level (left) and f_0 span (right) in literal and sarcastic conditions.

3.3. Duration

Duration (in s) was computed for each utterance as a whole. Table 2 provides means and standard deviations (in parentheses) for the duration of literal and sarcastic utterances, in all validated pairs. Lengthening was computed for each pair of utterances as the difference in duration between the sarcastic and the literal versions divided by the duration of the literal version. Mean lengthening from literal to sarcastic utterances was 29.79% (standard deviation 0.22). A paired-samples t-test was conducted to compare duration in literal and sarcastic conditions. There was a significant difference in duration for literal ($M= 1.295$, $SD=0.26$) and sarcastic ($M=1.667$, $SD=0.39$) conditions; $t(50)=-8.5249$, $p < 0.001$. These results suggest that sarcasm does have an effect on utterance duration. Specifically, when producing a sarcastic utterance, the utterance is approximately one third longer relative to the literal version.

Table 3. Mean and standard deviation (in parentheses) of utterance duration in literal and sarcastic conditions

	Duration	Lengthening
Literal	1.295 (0.26)	
Sarcastic	1.667* (0.39)	29.79% (0.22)

4. Discussion

Our results suggest that sarcastic intent can be recovered from an utterance, in absence of an explicit context. Specifically, the average accuracy score for all of the 468 utterances produced by 10 speakers was 79%. In order to evaluate which acoustic features signal sarcastic speech, we only conducted acoustic analyses on utterances which were robustly identified by a set of participants. 22% of the original literal/sarcastic pairs were identified as being literal/sarcastic at a confidence level of at least 4 (on a scale from 1 to 5) and were therefore retained. The acoustic analyses performed on the perceptually validated utterances show that sarcastic utterances display a higher pitch level (0.93 semitones) a wider span (2.39 semitones), as well as a longer duration (around one third longer).

Our study suggests that French speakers seem to use a higher pitch level and a wider pitch span to express sarcasm or irony. Note that our results are in line with those presented by [18] for French, in which higher pitch level was also reported for ironic utterances. Other authors have found a higher pitch level in sarcastic tone [23] for English. On the other hand, our results contrast with those reported for German [25] and English [3, 7, 10], in which lower instead of higher mean f_0 is reported for irony/sarcasm. A flat pitch has also been cited as an acoustic cue of sarcasm [4, 14, 15, 19, 21, 24].

Conflicting results are also reported concerning pitch span. Cheang and Pell [10] found a smaller f_0 range for sarcastic speech, while Fónagy [13] described several stages in the expression of irony, with first a chest voice and a creak effect, then a head voice with a rise in pitch and finally a chest voice with a low steady pitch. Several authors have found exaggerated pitch accents over the entire utterance, on all content words [4, 15, 26].

Note also that the specific class of ironic speech as well as language-specific implementation of sarcasm employed in previous studies might explain some of the differences found in the literature. Also, differences in the intonational phonology of each language, which might privilege either rising or falling pitch accents, might be a source of the crosslinguistic variability reported in the literature. More data (on both the acoustic/phonetic and phonological level) are hence needed in order to better account for our results and to determine whether level and span are sufficient cues to sarcasm or whether specific intonational contours (with pitch accents at specific locations and/or specific boundary tones) are also needed.

5. Conclusions

Our study shows that French sarcastic speech is produced through acoustic features that can be correctly identified even in absence of linguistic contextual cues. Specifically, sarcasm appears to be implemented through both heightened pitch level and pitch range expansion. Our results confirm previous findings reported in the literature on the production of irony in French, though being in contrast with some findings for Germanic languages, in which irony appears to be signaled through both pitch level lowering and pitch range compression. More cross-linguistic results are needed in order to confirm language-specific characteristics of sarcastic tone of voice.

6. Acknowledgements

We thank G.Bailly for his help with the phonetic transcription, F.Cangemi for a Praat script, T.Hueber and X. Laval for their

help during the production and perception experiments. We are sincerely grateful to all the participants.

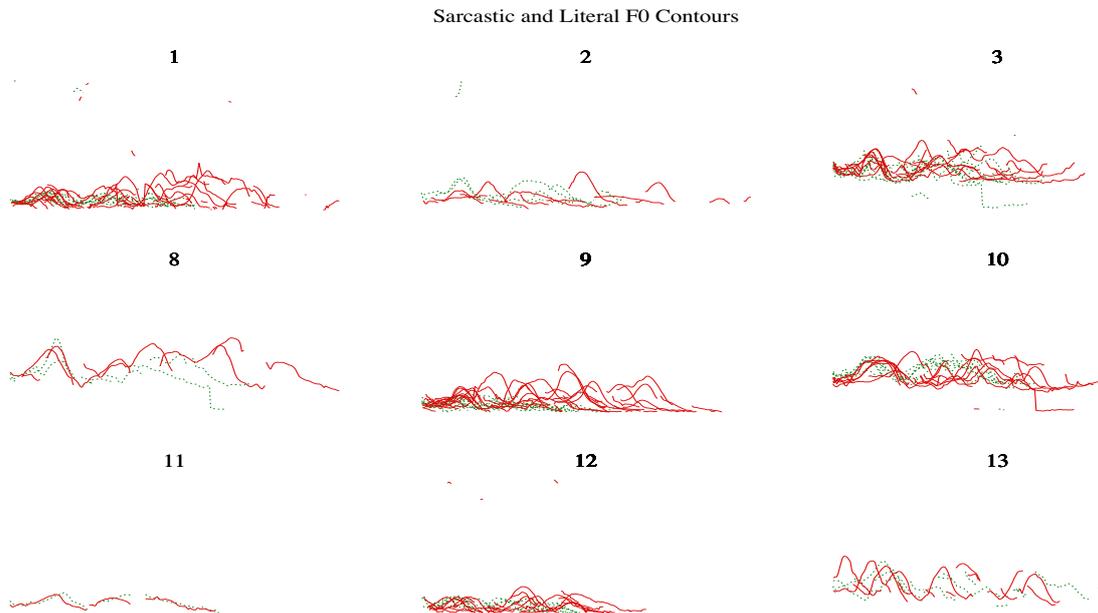


Figure 1: Superposition of literal (dashed green lines) and sarcastic (solid red lines) f_0 contours for each validated utterances for the 9 validated speakers. Sarcastic contours are longer and display a higher level and wider span.

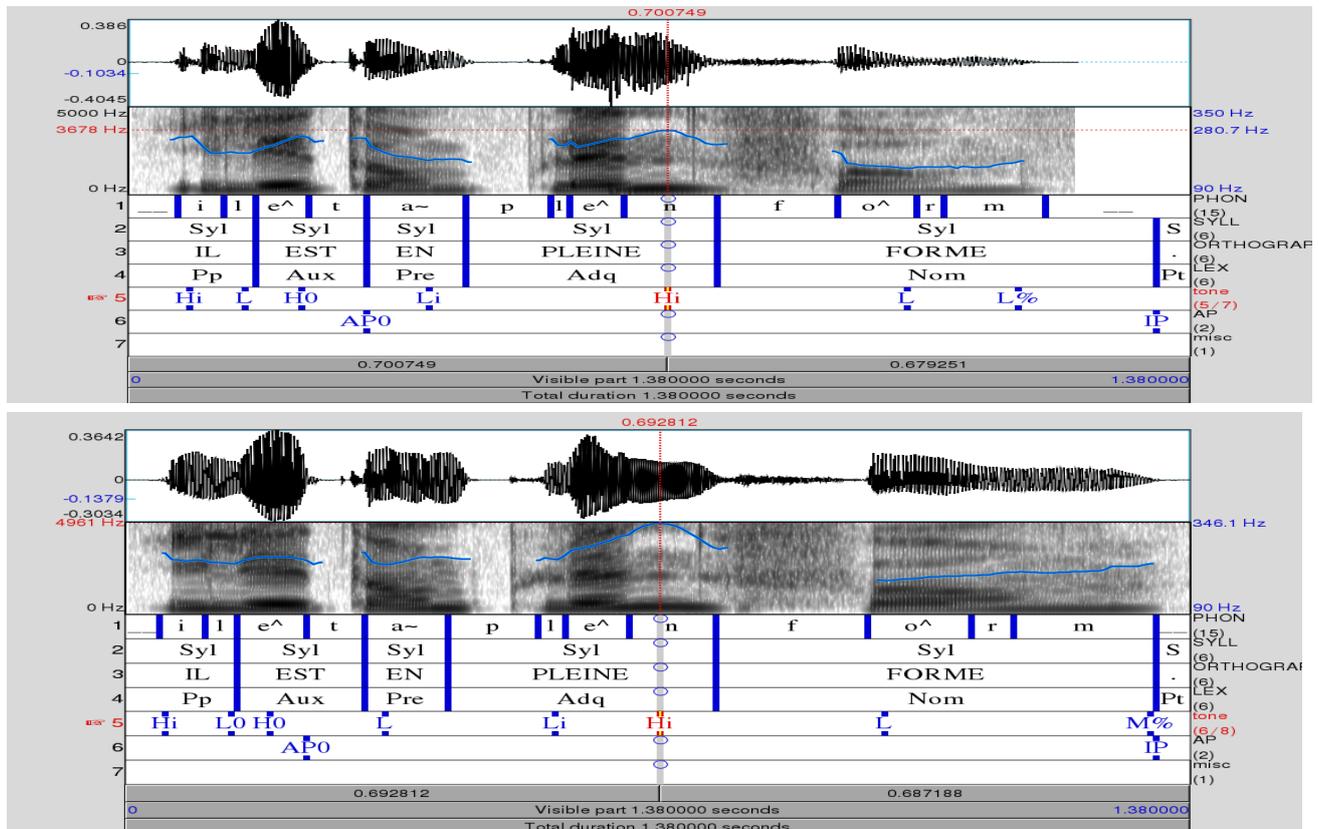


Figure 2: Production of literal (top) and sarcastic (bottom) versions of the sentence "Il est en pleine forme" (He is in great shape) by a female speaker. Maximum f_0 reaches 281 Hz in the literal version and 346 Hz in the sarcastic version.

7. References

- [1] Ackerman, B. "Young children's understanding of a speaker's intentional use of a false utterance". *Developmental Psychology*, 17: 472-480, 1981.
- [2] Ackerman, B. "Form and function in children's understanding of ironic utterances". *Journal of Experimental Child Psychology*, 35: 487-508, 1983.
- [3] Anolli, L., Ciceri, R. and Infantino, M.G. "Irony as a game of implicitness: Acoustic profiles of ironic communication". *Journal of Psycholinguistic Research*, 29: 275-311, 2000.
- [4] Attardo, S. Eisterhold, J., Hay, J. and Poggi, I. "Multimodal markers of irony and sarcasm". *Humor: International Journal of Humor Research*, 16(2): 243-260, 2003.
- [5] Boersma, P. and Weenink, D. "Praat: doing phonetics by computer" (Version 5.0.02) [Computer program]. Available from: <http://www.praat.org/>, retrieved 27 December 2007.
- [6] Bryant, G. and Fox Tree, J. "Recognizing Verbal Irony in Spontaneous Speech". *Metaphor and Symbol*, 17(2): 99-117, 2002.
- [7] Bryant, G. and Fox Tree, J. "Is there an Ironic Tone of Voice?" *Language and speech*, 48: 257-277, 2005.
- [8] Capelli, C. A., Nakagawa, N. and Madden, C. M. "How children understand sarcasm: the role of context and intonation". *Child Development*, 61, 1824-1841, 1990.
- [9] Champagne-Lavau, M., Charest, A., Anselmo, K., Blouin, G. and Rodriguez, J.P. "Theory of mind and context processing in schizophrenia: the role of cognitive flexibility". *Psychiatry Research*, 200: 184-192, 2012.
- [10] Cheang, H.S. and Pell, M. D. "The sound of sarcasm". *Speech Communication*, 50(5): 366-381, 2008.
- [11] Cutler, A. "On saying what you mean without meaning what you say", in M.W. LaGaly, R.A., Fox, and A., Bruck [Eds], *Chicago Linguistic Society*, Chicago, 117-127, 1974.
- [12] Cutler, A. "Beyond parsing and lexical look-up: an enriched description of auditory sentence comprehension". *New approaches to language mechanisms: A collection of psycholinguistic studies*: 133-149, 1976.
- [13] Fónagy, I. "Synthèse de l'ironie. Analyse par la synthèse de l'intonation motivée". *Phonetica* 23(1) : 42-51, 1971.
- [14] Fónagy, I. "La mimique buccale. Aspect radiologique de la vive voix. Radiological aspects of emotive speech". *Phonetica*, 33: 31-44, 1976.
- [15] Haiman, J. "Talk is cheap: Sarcasm, alienation, and the evolution of language". USA: Oxford University Press, 1998.
- [16] Ivanko, S.L. and Pexman, P. M. "Context Incongruity and Irony Processing". *Discourse Processes*, 35(3): 241-279, 2003.
- [17] Ladd, D. R.. "Intonational Phonology". Cambridge: Cambridge University Press, 1996/2008.
- [18] Laval, V. and Bert-Erboul, A. "French-speaking children's understanding of sarcasm: The role of intonation and context". *Journal of Speech Language and Hearing Research*, 48(3): 610-620, 2005.
- [19] Milosky, L. and Wroblecki C. A. "The Prosody of irony". Paper presented at the International Society for Humor Studies Conference, Ithaca, NY. 1994.
- [20] Morlec, Y., Bailly, G. and Aubergé, V. "Generating prosodic attitudes in French: Data, model and evaluation". *Speech Communication*, 33: 357-371, 2001.
- [21] Myers Roy, A. "Towards a definition of irony", in Fasold, Ralph W. and Roger Shuy [Eds], *Studies in Language Variation*, 171-183, University Press, 1976.
- [22] Pexman, P. "It's Fascinating Research. The Cognition of Verbal Irony". *Current Directions in Psychological Science*, 17: 286-290, 2008.
- [23] Rockwell, P. "Lower, slower, louder: vocal cues of sarcasm". *Journal of Psycholinguistic Research*, 29(5): 483-495, 2000.
- [24] Shapely, M. "Prosodic variation and audience response". *IPrA: Papers in Pragmatics*, 1(2): 66-79, 1987.
- [25] Sharrer, L. and Christman, U. "Voice Modulations in German Ironic Speech". *Language & Speech*, 54(4): 435-465, 2011.
- [26] Uhmans, S. "On rhythm in everyday German conversation: Beat clashes in assessment utterances", In E. Couper-Kuhlen and M. Selting [Eds], *Prosody in Conversation: Interactional Studies*, 303-365, Cambridge University Press 1996.