

Angelika Braun and Astrid Schmiedel

The phonetics of ambiguity: A study on verbal irony

Abstract: Wordplay and verbal irony are not as different as they might seem at first glance. In fact, they share the property of being *uneigentlich* ('non-actual'), i.e., the meaning of an utterance is somehow dissociated from its wording. Thus the question arises whether this kind of speech is signaled in the production process, and whether the cues thus created can be detected by naive listeners. This contribution presents an overview of the various means by which verbal irony may be conveyed to the listener. This is followed by an empirical study which demonstrates that verbal irony is indeed inherently ambiguous and that the underlying message is phonetically coded. The question addressed is how disambiguation takes place on the phonetic level. In other words – how does a speaker signal the intended meaning and how are listeners able to get the underlying message. The study deals with single-word utterances in sincere and ironic settings. Parameters considered include average voice fundamental frequency (F0) and related measures, intensity of the voice signal, and duration. The results of the acoustic measurements show that production results vary with type of irony (sarcasm vs. kind irony). It is argued that the underlying emotional states have to be taken into account for the interpretation of the measurement results. The perception study yields overall recognition rates of about 70 %, the sincere utterances being identified significantly better than the sarcastic ones in the positive stimulus set (sarcasm) and the ironic ones better than the sincere ones in the negative stimulus set (kind irony).

Keywords: articulatory precision, emotions, formants, intensity, kind irony, loudness, paralinguistic cues, perception of verbal irony, phonetics, sarcasm, sincerity, speaking tempo, (verbal) irony, (voice) pitch, voice quality

1 Introduction

At first glance, verbal irony and wordplay do not seem to have much in common. However, in rare cases like the following one, they may actually coincide:

Mein Trumpf (banner at an anti-Trump demonstration; seen in the German TV-program *ZDF Morgenmagazin* on 11 November 2016)

This example may be interpreted in two different ways: If one focuses on the graphemic level, the German *Trumpf* [tʁʊmpf] ‘trump (card)’ can be regarded as playing on the name of *Trump* in a sarcastic manner. The personal possessive *mein* ‘my’, however, suggests a different interpretation: If *Trumpf* is pronounced [tʁʌmpf], *mein Trumpf* will rhyme – at least to German ears – with Hitler’s *Mein Kampf*. This would mean to draw an analogy between the two men. Thus, wordplay may in rare instances be sarcastic, but verbal irony normally will not take on the form of wordplay.

However, both wordplay and verbal irony share the property of being non-actual speech (*uneigentliches Sprechen*; cf. Berg 1978), i.e., the meaning is somehow dissociated from the wording of an utterance.¹ Just how this dissociation is implemented forms a question which is for phoneticians to answer.

Irony can be conceived as an extreme case of (pragmatic) ambiguity, as there are (usually two) potential interpretations of an identical wording which are mutually incompatible, and it is up to the listener² to disambiguate the utterance and choose the interpretation which is contextually appropriate. From a pragmatic perspective, the concept of irony relies on the listener being able to detect this divergence – otherwise its purpose would be lost. Verbal irony thus creates a kind of camaraderie between speaker and (part of the) listeners.³

Even though the above definition of verbal irony seems pretty straightforward at first glance, it is by no means as clear if one takes a closer look. For example, it may prove to be difficult to separate verbal irony from joking or figures of speech like metaphors (*Donald Trump really hit the spot with Kim Yong Un*) or litotes (*not bad* meaning ‘very good’). On a more principal strand, Gibbs and O’Brien (1991) point out that (a) verbal irony does not necessarily involve a difference between what is said and what is meant and that (b) verbal irony does not necessarily constitute an intentional speech act. An example for (a) would be as follows: Assume that someone broke a window in the Browns’ kitchen by kicking a football right into it. Mrs. Brown asks the neighbors’ son James whether he did it. He, who was actually the one who broke the window replies: “Sure I did!”

1 There are exceptions to this rule; cf. the observations on fake irony and unintentional irony below.

2 Since this paper is dealing primarily with verbal irony, it is mainly concerned with spoken utterances as opposed to written ones. That is why the term “listener” is used to denote the recipient of a message.

3 Instances of verbal irony at the expense of a person present but not aware of the ambiguity are conceivable. For example, if there are three people playing a card game and one plays the wrong card, one of the other two may say “great play”, with only the third player getting the irony and the one having played the card feeling flattered.

Objectively speaking, he is telling the truth, but by his tone of voice he may indicate that he is being sarcastic and therefore trying to convince Mrs. Brown that he does not mean what he is saying. This process could be termed *fake irony*. An example for (b) would be someone who has just been hit by a pigeon's droppings on his shoulder without being aware of it, saying: "I really don't see why people dislike the pigeons so much. They don't do any harm." In this case, everybody but the speaker is aware of the ironic nature of the utterance. It can therefore be called *unintentional irony*.

In view of examples like the above, it is hard to disagree with Haverkate (1990: 106), who concludes that "[...] it is difficult, if not impossible, to register, describe and explain in an exhaustive manner all aspects of that fascinating phenomenon that we call 'irony'".

There is a broad consensus in the literature that ironic speech is signaled by some kind of "tone of voice" (cf. e.g. Gibbs and Colston 2007; Cheang and Pell 2008). On the other hand, there are authors who argue "that there is no particular ironic tone of voice" (Bryant and Fox Tree 2005: 257; similarly Gibbs and O'Brien 1991; Winner et al. 1987). One of their arguments is that readers are able to detect irony in written language, which obviously does not contain any intonational cues. This can be countered by the fact that extensive context information was available to the readers. The experiment described in the present contribution, on the other hand, deals with very short utterances which are presented devoid of context. It therefore seems legitimate to study the influence of vocal factors contributing to the decoding of verbal irony nonetheless. Furthermore, frequently there is no visual channel available to aid with disambiguation, e.g. when talking over the telephone. In these cases, the oral channel is the only one by which to convey the relevant paralinguistic⁴ information.

In a study on speech production, Rockwell (2000) described a very clear pattern of signaling irony as being "Lower, slower, louder". Subsequent studies showed that things may not be quite so simple. One reason is that in early studies on verbal irony (cf. e.g. Schaffer 1982; Rockwell 2000; Bryant and Fox Tree 2002) so-called sincere utterances, which could also be called affirmative, were compared to ironic ones. Yet, Anolli, Ciceri, and Infantino (2000, 2001, 2002) were able to show that a third category should be introduced in order to serve as a valid reference, i.e. neutral utterances which are neither actively affirmative nor ironic.

⁴ By paralinguistic we understand information which is conveyed by „tone of voice“. This comprises the classical features of voice fundamental frequency and related parameters, intensity, and duration, as well as laryngeal and supralaryngeal voice quality features.

On a different strand, two kinds of verbal irony need to be distinguished: the “genuine” sarcastic irony which can be characterized as using “blame by praise” (*great* to express disapproval) on the one hand and the so-called “kind irony” which can be characterized as using “praise by blame” (*terrible* to express admiration) on the other. It can by no means be taken for granted that those two follow the same pattern.

Relatively little attention has been given to sociolinguistic properties of speaker and listener groups. A few studies focus on intercultural differences (Cheang and Pell 2009, 2011; Adachi 1996 for Japanese), demonstrating difficulties in identifying verbal irony in a different linguistic or cultural context. However, there are hardly any studies focusing on age effects in the encoding or decoding phase.⁵ Schmiedel (2017: 122) describes some difficulties on the part of her older listeners in particular, but cannot draw conclusions owing to the small number of subjects. Very few researchers have studied gender effects with respect to either production or perception. Chen and Boves (2018) observe gender effects on various pitch and duration parameters on the production side.

The difference between what is meant and what is said may be looked at from the perspective of the speaker as well as the listener. On the part of the speaker, the ironic, neutral or sincere character of an utterance is cued primarily through context, but also by way of extralinguistic⁶ signaling. This may occur in various ways: the literal “tongue in cheek” or a wink being examples on a visual strand.

So far, studies on verbal irony have focused on the distinction between sincere (and, possibly, neutral) stimuli as opposed to sarcastic ones (“blame by praise”). Listeners were asked to classify utterances as either ironic or sincere and have proven to be quite successful at that task. The so-called kind irony (“praise by blame”) has not nearly received as much attention.

2 Phonetic means of coding non-actual speech

The relationship between literal and actual meaning of an utterance in ironic speech needs to be communicated to the listener. This applies to “normal”

⁵ In the context of this contribution, “age effects” refer to an adult population. There are quite a few studies on developmental aspects of the decoding of irony in particular (e.g. Creusere 2000; Laval and Bert-Erboul 2005; Wilson 2013) but this aspect is beyond the scope of the present study.

⁶ By extralinguistic we understand non-verbal, non-vocal means of communication, i.e. facial expression and gestures.

instances of irony, in which there is a discrepancy between the two, as well as to fake irony and unintentional irony. In fake irony it is absolutely essential to use an ironic tone of voice because otherwise the utterance would be interpreted as what it is, i.e. truthful. Unintentional irony, on the other hand, is characterized by precisely the absence of cues to ironic speech.

If all interlocutors are physically present during the exchange, irony is often signaled by nonverbal means (“tongue in cheek”, winking, a shrewd smile, etc.). If, however, the exchange takes place without the visual channel being available, e.g. over the telephone, the signaling has to be achieved by vocal and / or verbal means alone. There is quite a range of phonetic mechanisms which may be utilized to signal non-actual speech. They will be discussed in the following sections. So far, not all of them have been studied in conjunction with verbal irony. Whenever findings are available from the literature on the subject, they will be cited.

2.1 Speaking fundamental frequency (‘pitch’) and related measures

The probably single most important measure in determining physiological stress (Hansen and Patil 2007), affective states (e.g. Banse and Scherer 1996; Braun and Heilmann 2012), deception (Anolli and Ciceri 1997), but also verbal irony is the average *speaking fundamental frequency* (FO) with *pitch* as its perceptual correlate.⁷ It describes the number of vocal fold vibrations in the larynx per time unit and is usually measured in Hertz (Hz).⁸ Fundamental frequency is generally averaged over an utterance, and its mean or median is established. Deviations from neutral speech are considered to be the consequence of a change in laryngeal muscle tone as a physiological correlate of the emotional state or stress condition.

⁷ The correlation between the acoustic and the perceptual parameters is fairly good but far from perfect. The main reason for this is that the human perceptual organ does not operate in a linear way either with respect to frequency or intensity (see 2.2 below). For instance, the sensitivity of the human ear is much higher in the low frequency range, which also contains the main area of speech, than in the high frequency range. In order to compensate for this it has become good practice to use the semitone scale as opposed to the Hertz scale in order to perceptually reflect voice pitch and its derivatives like standard deviation and range. A similar principle applies to intensity.

⁸ Especially if male and female speakers are to be compared, it may be useful to express the results in semitones rather than Hz, thus eliminating the effect of the physiological difference between male and female voices and facilitating a comparison between male and female voices.

Besides the *mean F0*, various distributional measures have been introduced in order to capture intonational detail. The most common ones are *standard deviation* and *range*, the former reflecting the degree of monotony or melodiousness, the latter representing the highest and the lowest note produced. Some studies focus on more elaborate prosodic cues like the pitch contour or the timing of the final fall (Chen and Boves 2018).

With respect to verbal irony, previous studies have rendered varying results. Whereas some researchers find a generally lower mean F0 in sarcastic stimuli as opposed to sincere ones (cf. Fónagy 1971; Rockwell 2000; Anolli, Infantino, and Ciceri 2002; Cheang and Pell 2008; Nauke and Braun 2011; Scharrer, Christmann, and Knoll 2011), others (Schaffer 1981; Anolli, Ciceri, and Infantino 2000⁹; Bryant and Fox Tree 2005) establish higher values. When looking at the standard deviations, a similar picture arises: Schaffer (1981) describes higher values for the sarcastic stimuli, whereas the majority of studies report monotony in sarcastic as opposed to sincere speech (Fónagy 1971; Anolli, Infantino, and Ciceri 2002, Attardo et al. 2003, Cheang and Pell 2008; Nauke and Braun 2011; Chen and Boves 2018). Schaffer (1981) and Bryant and Fox Tree (2005) present opposite or inconclusive results. The findings for F0 range are just as contradictory: an increase in F0 range in sarcastic utterances was found by Schaffer (1981) and – to a limited extent – Bryant and Fox Tree (2005); a decrease by Anolli, Infantino, and Ciceri (2002), Cheang and Pell (2008) and Nauke and Braun (2011).

One can speculate about the reasons for these discrepancies. Scharrer, Christmann, and Knoll (2011: 3–4) largely attribute them to different languages being studied. However, it also seems worth considering that methodological issues may have played a role (e.g. the inclusion of different kinds of irony; length of stimuli; inclusion of neutral stimuli).

2.2 Vocal intensity ('loudness')

The sound pressure level or intensity of speech sounds with *loudness* as its auditory correlate¹⁰ is largely caused by an increase in subglottal pressure which is in turn influenced by the respiratory musculature. It can be measured in terms of RMS (root mean square) including standard deviation and range and is expressed in decibels (dB).

⁹ Experiment 1.

¹⁰ It should be kept in mind that the human ear is most sensitive to intensity differences within the frequency range of speech sounds.

Studies looking at verbal irony have come to different conclusions regarding the intensity measures. Rockwell (2000) finds an increase in average intensity in sarcastic stimuli, whereas Nauke and Braun (2011) come to the opposite conclusion. The latter finding corresponds to the results reported by Bryant and Fox Tree (2005), who established a lower intensity standard deviation in their sarcastic stimuli as opposed to sincere ones.

Probably it may prove worthwhile to look at the intensity dynamics of the stressed vowel in greater detail in the future. It is of interest whether the timing of the maximum intensity within the stressed vowel differs between ironic and sincere or neutral conditions or will even vary with type of irony. Schmiedel (2017) finds a highly significant delay of the intensity maxima in sarcastic, kindly ironic and sincerely negative stimuli as opposed to neutral utterances. No difference can be established between sincere praise and neutral utterances. This preliminary result demonstrates that this parameter deserves more attention.

2.3 Spectral tilt / center of gravity

The measurement of spectral tilt or the spectral center of gravity can be considered as an indication of the distribution of intensity over the range of frequencies across the spectrum. A predominance of lower frequencies corresponds to a muffled, soft voice, an emphasis of higher frequencies will result in what is perceived as an inherently loud, strong voice. The spectral composition of harmonics reflects the shape of the glottal impulse – the longer the closing phase of the vocal folds, the more harmonics will be visible in the spectrogram, the higher is the center of gravity, and the smaller is the spectral tilt. The present authors are not aware of any studies on verbal irony up to now which have used this parameter.

2.4 Tempo and pausing

Speaking tempo and *pausing* can be considered as interrelated to some extent, because the more pauses occur, the slower the tempo will be. It is advisable, though, to keep those two factors separate by differentiating between syllable rate and articulation rate. The former includes pauses whereas the latter does not, thus taking only net speech into account. Both syllable and articulation rate can be measured in different ways, depending on the linguistic unit per time unit calculated. The linguistic units most commonly used are sounds, syllables or words; the time units are seconds or minutes. The most widespread units are syl-

lables per second – this will, however, become problematic when comparing languages with largely different phonotactics. In those cases, the use of sounds per second may be more advisable.

A follow-up decision concerns the choice between phonetic and linguistic syllables. As an example, the realization of the German lexeme *haben* ‘to have’ as [ham] contains two linguistic syllables as opposed to one phonetic syllable. Calculating speaking tempo in terms of phonetic syllables will thus provide a close approximation of actual articulator movement whereas linguistic syllables will reflect the degree of articulatory precision (see 2.6 below). In short utterances, sounds per second may be the method of choice in studying speaking tempo.

There is a rare degree of consensus in previous studies in citing a decrease in speaking tempo in sarcastic stimuli as opposed to sincere ones (Fónagy 1971; Schaffer 1982; Rockwell 2000, Anolli, Infantino, and Ciceri 2002; Cheang and Pell 2008; Scharrer, Christmann, and Knoll 2011; Rao 2013; Chen and Boves 2018).

Pauses may take on both the form of filled and unfilled pauses. Fillers have been found to be a stable speaker specific feature (Braun and Rosin 2015). However, studying pauses only makes sense in longer utterances.

Studies which examine the number of pauses and pause length are very rare. Anolli, Infantino, and Ciceri (2002) report a larger number of pauses and at the same time significantly shorter pauses in ironic as compared to normal speech.

2.5 Articulatory precision

Articulatory Precision denotes the degree to which articulatory targets are reached. This presupposes that speaking entails constant movement of articulators from one target point to the next as opposed to producing a sequence of stationary segments as is suggested by alphabetical renditions of speech. In other words, the speech organs are in motion most of the time, whereas the stationary phases are practically non-existent. However, there are differences between various speaking styles with respect to articulatory precision. In careful speech, the targets are generally reached, whereas this is the case to a much lesser degree in everyday speech.¹¹ For instance, a plosive like /b/ may be pronounced as a fricative [β] in a word like *habe* ‘have’. In vowels, lack of articulatory precision means centralization, which implies a movement of the articulators towards the central vowel *schwa* [ə]. An example is the pronunciation of German *gut* ‘good’ as [gʊt^h] as opposed to [gu:t^h]. One way of determining the degree of articulatory precision

¹¹ Speaking while intoxicated will, however, dramatically reduce articulatory precision.

is to measure vowel formants. The more “peripheral” the formants are within the vowel trapezoid, the more precisely those vowels are realized. Precise articulation takes more time than sloppy one (Hildebrandt 1961). This is exemplified by the above *gut* example, where the more centralized vowel is also shorter than the more peripheral one. This is part of a general tendency: slower speaking tempo and higher articulatory precision will often co-occur.¹²

Different procedures have been put forward to measure articulatory precision. One of them is the *Lautminderungsquotient* ‘sound elision quotient’ (Hildebrandt 1961) where

$$LMQ = 10 - \frac{10 \times n_{\text{marked}}}{n_{\text{unmarked}}}$$

n_{marked} represents the number of phones realized in the marked condition (sample under investigation); n_{unmarked} represents the number of phones realized in the unmarked condition (reference / standard).

Positive values represent elisions (*wir ham* [ham] as opposed to *wir haben* ['ha:bən] ‘we have’); negative values represent epenthesis (German *faul* [fəʊl] ‘lazy’ being pronounced as ['fəʊəl]).

On a different strand, Low and Grabe (1995) developed the Paired Variability Index (PVI), which was originally intended to provide a tool which would allow to gradually distinguish between so-called syllable-timed and stress-timed languages. This measure might also be used to distinguish ludic from “unmarked” speech. Neither LMQ nor PVI seem to have been applied as yet in studies on verbal irony.

Scharrer, Christmann, and Knoll (2011) approach the issue of articulatory precision from an acoustic point of view. They measure formant frequencies and find a larger vowel space in ironic speech than in literal utterances. They take this to imply that there is vowel hyperarticulation in sarcastic speech (Scharrer, Christmann, and Knoll 2011: 19).

¹² This, once again, only holds true for “normal” speech. On the other hand, a reduction of speaking tempo may well be accompanied by lack of articulatory precision if the speaker is inebriated, whereas in emotional speech (specifically: anger), high articulatory precision may be accompanied by high tempo (cf. Kienast 2002: 81, 105).

2.6 Voice quality

The term “voice quality” refers to a number of laryngeal as well as supralaryngeal settings which contribute to the individual sound of a particular voice. Examples are the former US Secretary of State Henry Kissinger, who constantly used what is called *creaky voice* or the CNN anchor Richard Quest who frequently uses so-called *ventricular voice*, i.e. phonation of the false vocal folds in addition to or instead of the real ones. The former results in a very low pitch; consequently the individual vibrations of the vocal cords can be perceived. The latter causes a raspy, strained sounding voice. An example of supralaryngeal settings can be observed in the current US-President Donald Trump, who regularly articulates with protruded lips.

While the use of a certain voice quality may have become associated with certain speakers, this does not mean to imply that voice quality will not change with communicative setting. For instance, certain voice qualities have been found to be associated with different emotions, e.g. breathy voice is frequently observed in sad and fearful utterances (Braun and Heilmann 2012; Probst and Braun 2016).

Cutler (1974: 117) and Haverkate (1990: 80) describe ironic speech as nasalized. Fónagy (1971) and Schaffer (1981) find a high proportion of creaky voice in ironic speech. There is certainly much room for further systematic research beyond these isolated observations.

Cheang and Pell (2008) study voice quality from a strictly acoustic perspective and observe a lowered harmonics-to-noise ratio¹³ in their sarcastic stimuli.

3 A study on short utterances

While the phonetic dimensions of verbal irony are generally fairly well researched (see section 2 above), there are two main shortcomings which need to be addressed. Firstly, in most studies on verbal irony so-called sincere utterances, which have also been called affirmative or literal, are compared to ironic ones (e.g. Cutler 1974; Schaffer 1982; Rockwell 2000). Yet, Anolli, Ciceri, and Infantino (2000, 2001, 2002) are able to show that a third category should be introduced in order to serve as a valid reference, i.e. neutral utterances which are

¹³ Harmonics-to-noise ratio (HNR) describes the difference in energy between harmonic and noisy components of speech. Low HNR levels are perceived as rough or raspy; high HNR levels are perceived as clear (Baken and Orlikoff 2000: 282).

neither actively affirmative nor sarcastic. These authors find both types of irony to deviate from neutral stimuli in the same direction. For instance, mean F0 as well as its range and standard deviation are higher in their ironic stimuli than in the neutral ones. The same applies to speech intensity. Finally, while the number of pauses in ironic speech exceeds that in neutral speech, pause length is shorter and articulation rate slower.

Secondly, most previous work on ironic speech is focused on sarcasm, i.e. blame by praise. Kind irony has not nearly received a comparable amount of attention. Furthermore, research has so far relied on utterances of sentence or phraselength (cf. e.g. Rockwell 2000; Anolli, Ciceri, and Infantino 2000; Cheang and Pell 2008; Scharrer, Christmann, and Knoll 2011).

The experiment presented in this contribution¹⁴ pursues a different path: It seeks the ultimate challenge in that it aims to determine whether speakers are able to signal sarcasm and kind irony in single-word utterances consisting of a maximum of two syllables in such a way that listeners can reliably decode them. This experimental approach was chosen with a forensic perspective in mind. In the forensic context, a defendant might argue that the *great* which he or she uttered in reply to a proposal by a hired killer to do away with his / her spouse on that same day was not sincere and therefore cannot be used as proof of guilt. In cases like that, a phonetician may be asked to testify in court whether a judgement on the sincere or sarcastic nature of an utterance is at all possible if it is based on a single syllable only.

This study furthermore draws a distinction between sarcasm (“blame by praise”) and kind irony (“praise by blame”) and includes neutral utterances as a baseline. The following research questions are asked:

- Do even very short sincere and ironic utterances differ with respect to their phonetic properties?
- Are there differences with respect to the type of irony (sarcasm vs. kind irony)?
- Are there gender differences in the production of irony in such very short utterances?
- Are listeners able to distinguish between very short sincere and ironic utterances?
- Are there gender differences in the perception of irony?

¹⁴ The results reported here consist in a subset of those contained in Schmiedel (2017). They are interpreted in a wider context here.

3.1 Materials and methods

When describing the experiment, we adopt the following terminological conventions: “Irony” is used as a cover term for both *sarcasm* and *kind irony*. “Sarcasm” denotes the *blame by praise* type of irony, i.e. *great* as a reaction to a missed train connection. “Kind irony” is used as a realization of irony through *praise by blame*, i.e. *terrible* as a reaction to a very good but less than perfect grade. The terms “sincere” or “affirmative” refer to the opposite of either sarcasm or kind irony, denoting praise in the former and blame in the latter case. They are distinct from “neutral” utterances which are unmarked.

3.1.1 Production experiment

A total of 20 speakers were recorded. They were balanced for sex (10 men, 10 women) with an average age of 27 years. Based on short scripted scenarios, they each produced 20 utterances consisting of monosyllabic (N = 7) or disyllabic (N = 13) words both in a sincere and an ironic context, 10 of which fell into the category of sarcasm (“blame by praise”) and 10 into the category of kind irony (“praise by blame”). In addition, the stimuli were recorded in a neutral setting, i.e. in read speech. In order to do that, the utterances were randomized with three repetitions each and then turned into a PowerPoint presentation where each stimulus appeared for three seconds. This was presented to the speakers without context. This procedure was chosen in order to avoid the typical “list intonation”. Table 1 shows the lexemes and their translation into English, Table 2 contains examples of scenarios.

Tab. 1: Positive (sarcasm) and negative (kind irony) stimulus sets (Translation: AB)

posi- tive	<i>danke</i> 'thanks'	<i>klar</i> 'course'	<i>klasse</i> 'great'	<i>lecker</i> 'tasty'	<i>nett</i> 'nice'
	<i>schön</i> 'good'	<i>Spitze</i> 'excellent'	<i>super</i> 'super'	<i>toll</i> 'awesome'	<i>Wahnsinn</i> 'phenomenal'
nega- tive	<i>Blödmann</i> 'dumbass'	<i>mies</i> 'lousy'	<i>Mist</i> 'shoot'	<i>schade</i> 'too bad'	<i>Schande</i> 'shame on you'
	<i>schlimm</i> 'bad'	<i>schrecklich</i> 'terrible'	<i>Streber</i> 'smart-ass'	<i>übel</i> 'nasty'	<i>Verdammt</i> 'damn it'

Tab. 2: Sample scenarios for sarcasm and kind irony (translation: AB)

Utterance	Scenario sincere	Scenario ironic
	Paul and Paula are going out to dinner. They are looking at the menu.	Paul and Paula are going out to dinner. They are looking at the menu.
<i>lecker</i> 'tasty'	Paul: Look, they've got your favorite pasta! Paula: Tasty!	Paul: Look, they've got frog's legs and snails. How does that sound? Paula: Tasty!
<i>Blödmann</i> 'dumbass'	It is Paula's birthday, and so far, her boyfriend Paul has neither congratulated her nor given her a present. He calls in the evening. Paul: Hi, I'm sorry. I've gotta work late. Don't wait for me. Paula: You do know what day it is today? Paul: Yes, it's Wednesday, why? Paula: Dumbass!	It is Paula's birthday, and so far, her boyfriend Paul has neither congratulated her nor given her a present. She comes home disappointed. When she enters the apartment, she finds that Paul and their friends have prepared a surprise party. Paul (smiling): I'm sure you thought I'd forgotten! Paula: Dumbass!

Measurements were carried out using the *praat* software package, version 5.3.04 (Boersma and Weenink 2012) with respect to duration, fundamental frequency, and intensity of the stimuli. The SPSS software package version 20.0 was used for the statistical analyses. These included *t*-tests and one-way repeated measures ANOVA, depending on the nature of the hypothesis to be tested.

3.1.2 Perception experiment

In order to test whether listeners are able to detect irony from short speech stimuli alone, a total of 44 listeners were studied. 23 of them women, 21 men. They were 28 years old on average. Thus, speakers and listeners were almost exactly of the same age.¹⁵ Only ironic and sincere utterances were used in the listening experiment. The neutral stimuli were not included in order to limit the duration of the experiment.

¹⁵ This is worth mentioning because it has been shown in relation to other perceptual tasks like age estimation that listener performance varies with the difference in age between speakers and listeners (Shipp and Hollien 1969).

Listeners were presented with a total of 800 stimuli (20 speakers x 20 words x 2 conditions (sincere and ironic)). Each stimulus was presented twice, separated by a 500 ms pause. The pause between different stimuli was set to 2 seconds. After every tenth stimulus, a 250-Hz tone was inserted in order to enable listeners to keep track of the utterances.

Samples were presented in isolation, i.e., listeners did not receive the scenarios. They had to tick one of two boxes labelled *ironic* or *sincere*. The duration of the listening experiment amounted to 50 minutes in total, which was judged to be too demanding to be carried out in one session. It was therefore split up into three chunks of two times 300 and once 200 stimuli which were judged in separate sessions. The pause between sessions lasted five minutes. Six utterances which were not part of the experiment were used for training purposes.

3.2 Results

When reporting the results, we adopt the following terminological conventions: “irony” is used as a cover term for both sarcasm and kind irony. The terms “sincere” or “affirmative” refer to the opposite of either sarcasm or kind irony, denoting praise in the former and blame in the latter case. They are distinct from “neutral” utterances which are unmarked. We furthermore distinguish between the “positive” and the “negative” stimulus sets. The former refers to the semantically positive lexemes, i.e. *toll* ‘great’ or *nett* ‘nice’. When used in a sincere manner, these utterances constitute praise, when they are used ironically, they constitute sarcasm. The latter, on the other hand, comprises semantically negative lexemes like *Mist* ‘shoot’ or *verdammt* ‘damn’, thus denoting either kind irony or blame.

3.2.1 Production experiment

3.2.1.1 Duration

Figure 1 shows the results for the timing parameters in the different scenarios as well as the neutral condition. In single-syllable utterances, there is no point in computing syllable rate or articulation rate. Measurements are thus confined to utterance length. Due to the difference in length between the sarcastic and kindly ironic stimuli, the absolute durations of both cannot be compared. Instead, the differences within each stimulus set will be looked at. As far as the positive stimulus set is concerned, it emerges that the sarcastic utterances are considerably longer than those expressing praise, the latter being only marginally longer

than the neutral stimuli. The same is true in principle for the negative stimulus set, but the difference between kind irony and blame is much smaller in this case. Both are markedly faster than the neutral stimuli. Statistical analyses were carried out in order to test the significance of these results. The difference in duration is significant for both types of irony (Wilcoxon two-tailed *t*-Test), however, results are much clearer for sarcasm than for kind irony. This is consistent with most previous research which was carried out on longer utterances. Specifically, Anolli, Ciceri, and Infantino (2000: 287) report a higher rate of articulation for praise as opposed to sarcastic irony and for blame as compared with kind irony. Their “normal speech” is fastest in both cases (cf. also section 2.4 above for further findings by previous authors).

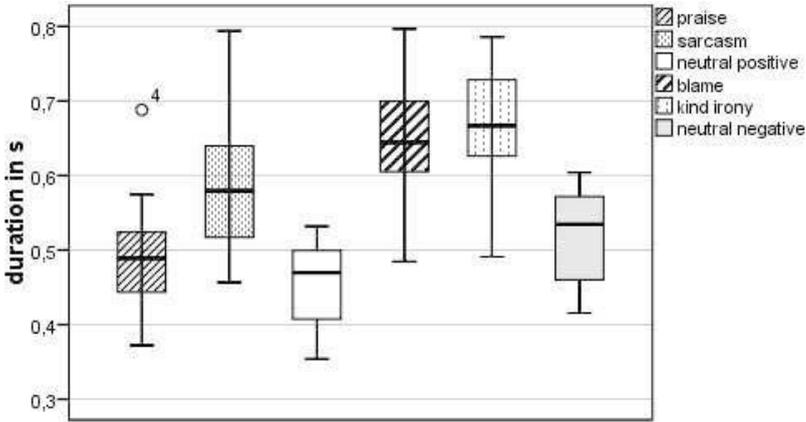


Fig. 1: Utterance duration in the various scenarios as well as the neutral condition

3.2.1.2 Fundamental frequency

The results on fundamental frequency parameters are summarized in Table 3. They include the mean as an indication of voice pitch and the standard deviation and range as indicators of melodiousness. Both standard deviation and range are expressed in terms of semitones in order to keep the absolute values for male and female voices comparable.

Tab. 3: Fundamental frequency: Mean, standard deviation, and range

Speakers	Stimulus set	Stimulus type	F0 mean (Hz)	SD (ST)	range (ST)
Female Ss.	positive set	praise	247	3.89	10.95
		sarcasm	194	2.45	7.54
		neutral	209	1.65	5.40
Female Ss.	negative set	blame	206	2.84	8.71
		kind irony	210	3.34	10.39
		neutral	209	1.54	5.36
Male Ss.	positive set	praise	149	3.65	11.31
		sarcasm	127	2.44	8.20
		neutral	125	1.70	5.85
Male Ss.	negative set	blame	130	2.49	8.31
		kind irony	137	3.20	9.86
		neutral	125	1.55	5.57

What has to be noted first of all is that there is a striking similarity between the neutral stimuli in both conditions. In fact, the difference between the two proved not to be significant for any of the F0 parameters. We thus join Anolli, Ciceri, and Infantino (2000: 290) in taking this as “an important indicator that confirms the internal validity of the experiment”. It also underlines the necessity for neutral stimuli to be included as a baseline in contradistinction to “sincere” ones, and it confirms that our subjects did very well at keeping those utterances neutral indeed.

In the positive stimulus set, the sarcastic utterances show a markedly lower mean fundamental frequency than do the sincere ones. This applies to male and female speakers alike. Both differences are highly significant ($p = .000$). The neutral renditions are closer to the sarcastic ones with a tendency to be slightly higher, though, for the female speakers, whereas they do not seem to follow a clear pattern in our male speakers.

Kind irony is much less clearly marked by mean fundamental frequency than sarcasm. In most male and female speakers alike, however, ironic stimuli are higher in mean F0 than the sincere ones, but the difference is much smaller than it is in the positive stimulus set and fails to reach significance. The deviations from neutral are also fairly small.

The two stimulus sets also differ with respect to *standard deviation*. Whereas $F0_{SD}$ is much lower in sarcasm than for praise, it is considerably higher for kind irony as compared to blame. Only the differences in the positive stimulus sets are significant, though ($p = .005$ for female speakers and $p = .002$ for male speakers). Both the sincere and ironic stimuli show a much larger standard deviation than the neutral utterances (sign. $p = .000$).

When it comes to fundamental frequency *range*, the sarcastic stimuli in the positive stimulus set show a much smaller span than the sincere ones. This is true for male and female speakers alike (sign. $p = .000$ for both). The values for the neutral stimuli are still more monotonous in 14 out of 20 cases and about equal to the sarcastic ones in four.

Quite the opposite results are to be found in kind irony. Here, ironic utterances show a range equal to or larger than that of sincere samples (sign. $p = .023$ for the female speakers and $p = .037$ for the male speakers). The neutral stimuli display the smallest $F0$ range.

It emerges that all three $F0$ parameters considered here show opposing results depending on the kind of irony studied. Thus they seem to be less of a marker of “ironic speech” as such but rather one of attitude, a positive attitude being expressed by a higher mean and a larger degree of melodiousness. This applies to both female and male speakers, although it is more marked in the former group.

3.2.1.3 Intensity

Once again, the neutral stimuli are very similar for both stimulus sets and for male and female speakers alike. Neither difference was found to be significant (Schmiedel 2017: 88). Differences emerge, however, when irony comes into play: positive wordings, whether sarcastic or sincere, tend to be louder than the neutral renditions, whereas negative wordings, whether kindly ironic or sincere, tend to be softer than the neutral stimuli. In other words, there is no clear distinction between “ironic” and “sincere” stimuli as such, but instead semantically positive utterances are louder than semantically negative ones.

As far as the positive stimulus set is concerned, there is a clear distinction between ironic and sincere utterances in that the former show a much lesser intensity as compared to the latter. The opposite is true for the n: In this case the sincere stimuli are softer than the ironic ones, albeit to a lesser degree.

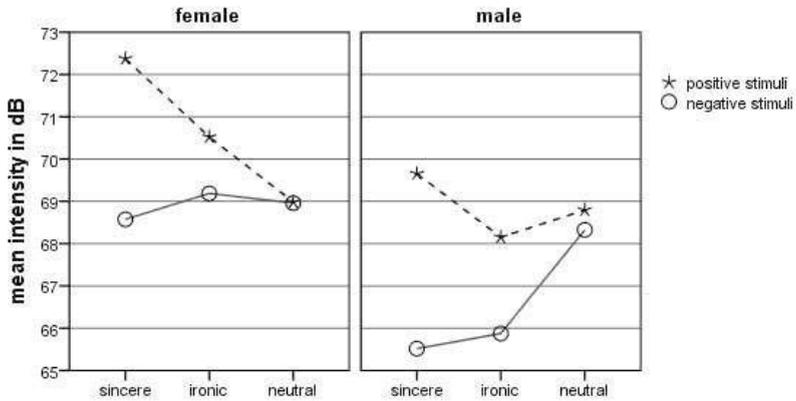


Fig. 2: Mean intensity in sincere, ironic, and neutral stimuli

Gender differences emerge with respect to intensity, not in kind, but in degree: female speakers tend to mark both kinds of irony by a greater intensity as compared to neutral utterances, whereas male speakers show a lower intensity.

These results are in keeping with those reported by Anolli, Infantino, and Ciceri (2002) with respect to the female speakers only; results for the male speakers point to the opposite direction. It seems that many of our male speakers are in the group which Anolli, Infantino, and Ciceri (2002) classify as group (4), for whom sarcasm is similar to cold anger and thus low in intensity.

Table 4 summarizes the findings on the production side including the significance levels established.

Tab. 4: Summary of findings – production

Feature	Type of irony	Relation	Sincere speech	Probability
Duration				
Mean	kind irony	>	blame	$p = .035^c$
	sarcasm	>	praise	$p = .000$
F0 women				
Mean	sarcasm	<	praise	$p = .000^a$
Mean	kind irony	>	praise	n.s.
SD	sarcasm	<	praise	$p = .005^b$
SD	kind irony	>	blame	n.s.
Range	sarcasm	<	praise	$p = .000^a$
Range	kind irony	>	blame	$p = .023^a$
F0 men				
Mean	sarcasm	<	praise	$p = .003^a$
Mean	kind irony	>	blame	n.s.
SD	sarcasm	<	praise	$p = .002^a$
SD	kind irony	>	blame	n.s.
Range	sarcasm	<	praise	$p = .000^a$
Range	kind irony	>	blame	$p = .037^a$
Intensity				
Mean	sarcasm	<	praise	$p = .002^a$
	kind irony	>	blame	n.s.

^a *t*-test paired; ^b Wilcoxon; ^c female speakers only (*t*-test paired)

3.2.2 Perception experiment

In the perception experiment, listeners had to decide in a forced-choice design whether a certain utterance was sincere or ironic. Given that the utterances they heard were disyllabic at the most, the task is much more difficult than in most, if not all, previous research. In view of this, recognition rates could be expected to be lower than those for longer utterances, where the paralinguistic information can be exploited to a much greater extent.

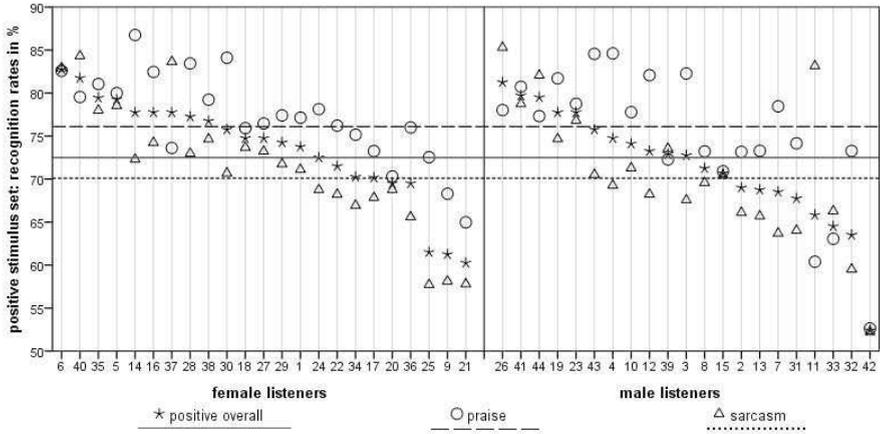


Fig. 3: Recognition rates for the “positive” stimulus set (sarcasm)

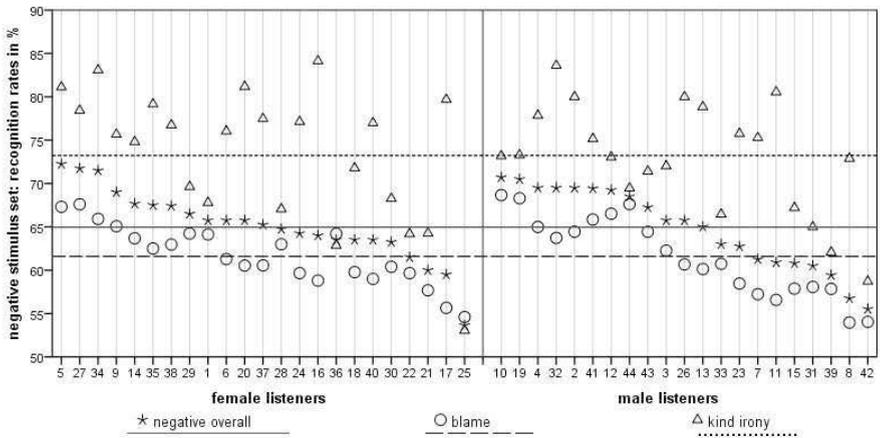


Fig. 4: Recognition rates for the “negative” stimulus set (kind irony)

The results are shown in Figures 3 and 4. When looking at them, it is immediately evident that the overall recognition rate for the positive (sarcastic) stimulus set is much higher than that for the kindly ironic stimulus set (72 % vs. 65 %). This, however, cannot be taken to imply that the kindly ironic utterances as such are identified less frequently than the sarcastic ones. On the contrary, the stimuli exhibiting kind irony were the ones which were recognized correctly more often

(73 %; cf. Figure 4) than the latter (70 %, cf. Figure 3). In fact, the overall recognition rate was best for praise (77 %), followed by kind irony and sarcasm, while blame was recognized only 62 % of the time. This means that the two kinds of irony are handled quite differently by the listeners: whereas the “praise by blame” type of irony is detected slightly better than the corresponding sincere utterances, the opposite is true for the “blame by praise” type of stimuli.

Female speakers’ utterances are recognized significantly (t-test paired; cf. Table 5) better than those produced by male speakers with the exception of blame. Listener sex, on the other hand, did not constitute a significant factor in the recognition rates.

Tab. 5: Summary of results of the perception experiment

Stimulus set			
positive (sarcasm)	>	negative (kind irony)	$p = .000^a$
Utterance type			
praise	>	sarcasm	$p = .000^a$
blame	<	kind irony	$p = .000^a$
Speaker sex			
praise / sarcasm / kind irony: female Ss	>	male Ss	$p = .000^a$
blame: female Ss	<	male Ss	$p = .000^a$

^a *t-test paired*

4 Discussion

It is not easy to compare the present findings to previous ones because of essential differences in the language studied and/or the experimental set-up. The latter concerns stimulus duration as well as the (lack of) inclusion of different kinds of irony and of neutral stimuli.

Firstly and most importantly, our results show that there evidently is a way of paralinguistic signaling of irony, which allows listeners to classify even very short context-free utterances as ironic or sincere at a level much beyond chance. It is, in fact, remarkable how well listeners are able to distinguish sincere from ironic monosyllabic words. From a forensic point of view, it is quite likely that listeners can reliably infer irony based on a single syllable even if no context

information is available. A major task for future work is thus to determine which of the acoustic parameters will best explain the perception results.

As far as production is concerned, the results for our positive stimulus set (a longer duration and a lower F0 in sarcasm than in sincere praise) are in good agreement with most previous studies which cover sarcasm only (cf. Cheang and Pell 2008; Rockwell 2000; Pexman and Olineck 2002; Scharrer, Christmann, and Knoll 2011; Rao 2013, to name only a few). They are clearly at variance, however, with Anolli, Ciceri, and Infantino's (2000) findings on fundamental frequency and also with other studies which cover languages other than German (Løevenbrück et al. 2013; Laval and Bert-Eboul 2005; Adachi 1996; Cheang and Pell 2009). This difference finds a plausible explanation in the fact that different languages were involved. Stimulus duration (monosyllables vs. full utterances) may also have played a role.

Our findings for the negative stimulus set (kind irony), particularly those with respect to F0 and intensity, form a mirror image those for the positive one, making it very clear that the two kinds of irony should not be confounded. With respect to the variables analyzed here, men and women pursue similar strategies in encoding the two kinds of irony. There are, however, differences in degree. This may have had consequences for their recognizability (see below).

Despite the clear distinction between sarcasm and kind irony, we do not unequivocally share Bryant and Fox Tree's (2005) conclusion that there is "very little support for the notion of an ironic tone of voice; that is, prosodic consistency across verbal irony utterances" (2005: 273). With respect to sarcasm, at least, typical patterns can be observed.

Whether these paralinguistic cues are exclusive to irony, however, remains open to question. One clear indication against a general classification of utterances as either "ironic" or "sincere" is the fact that the distinction between sarcasm and kind irony which has been pointed out by Anolli, Ciceri, and Infantino (2000) is supported by our results. All production parameters were found to be opposite for sarcasm and kind irony as compared to their sincere counterparts. This means that at least those two variants of irony should be taken into consideration in future research, because only the comparison between ironic and sincere stimuli of different sorts will reveal the different patterns. The same applies to the introduction of the "neutral" category, which may serve as some kind of a baseline.

Secondly, the markers of verbal irony as established in the present study do not seem to be confined to ironic use. As an example, we propose that for the F0 parameters it is the "underlying sentiment" that determines their relationship between the members of a stimulus pair as opposed to the "face value" of the

lexemes. We argue that this is the main reason why sarcasm and kind irony behave so differently in many respects while both are distinct from “neutral”. For instance, utterances like the sarcastic *toll* ‘great’ are longer, lower-pitched, and lower in intensity than the sincere *toll*, whereas the kindly ironic *super* ‘super’ is longer, higher pitched, and higher in intensity than the corresponding sincere token.

In fact, the findings for our positive stimulus set closely resemble the results reported by Braun and Heilmann (2012) on cold anger. Those authors found cold anger to exhibit a lower mean F0 and a smaller standard deviation and range than emotionally neutral utterances. At the same time, the patterns established for praise in the present study are very similar to those for joy as observed in Braun and Heilmann (2012). Our results strongly suggest a closer look at the relationship between emotional speech and ironic speech. This view corresponds perfectly with Anolli, Infantino, and Ciceri’s (2002) description of what they call a “family of ironic voices” (2002: 374). Our data for kind irony match their pattern 1a, which they characterize as “bantering joy”, and our sarcastic data clearly correspond to their pattern 2b (“cold anger”; Anolli, Infantino, and Ciceri 2002: 374). Bryant and Fox Tree (2005) argue along that same line when they conclude: “Speakers are communicating multiple messages by layering propositional and non-propositional information, and providing prosodic cues in contextually dependent ways that map differentially onto the simultaneously presented information [...]” (2005: 272). It will be a prime task for future research to separate and identify those layers of information and their phonetic correlates.

Thirdly, as far as perception is concerned, there are clearly individual speakers whose ironic signals are more easily decoded than those of others. Generally speaking, female speakers’ cues to irony were identified significantly better than those uttered by male speakers. This once again has a parallel in the processing of emotional speech. Braun and Heilmann (2012) found female listeners to be highly significantly better at recognizing emotions than male listeners.

On the other hand, there are listeners who are better at decoding speaker intent than others. This time, however, there was no clear distinction along gender lines. It would be of future interest to establish whether there is a correlation between the ability of a given listener to spot irony and his or her ability to show empathy. If those two abilities are found to be related, further studies on patients who are unable to express empathy or detect emotions, who e.g. have been diagnosed with Asperger autism, would have to constitute the next step.

5 References

- Adachi, Takanori. 1996. Sarcasm in Japanese. *Studies in Language* 20(1). 1–36.
- Anolli, Luigi & Rita Ciceri. 1997. The Voice of Deception: Vocal Strategies of Naïve and Able Liars. *Journal of Nonverbal Behavior* 21(4). 259–284.
- Anolli, Luigi, Rita Ciceri & Maria Giaele Infantino. 2000. Irony as a Game of Implicitness: Acoustic Profiles of Ironic Communication. *Journal of Psycholinguistic Research* 29(3). 275–311.
- Anolli, Luigi, Maria Giaele Infantino & Rita Ciceri. 2001. “You’re a Real Genius!”: Irony as a Miscommunication Design. In Luigi Anolli, Rita Ciceri & Giuseppe Riva (eds.), *Say not to Say: New perspectives on miscommunication. Bd. 3. Emerging Communication: Studies in New Technologies and Practices in Communication*. IOS Press, 141–164.
- Anolli, Luigi, Maria Giaele Infantino & Rita Ciceri. 2002. From “blame by praise” to “praise by blame”: Analysis of vocal patterns in ironic communication. *International Journal of Psychology* 37. 266–276.
- Attardo, Salvatore, Jodi Eisterhold, Jennifer Hay & Isabella Poggi. 2003. Multimodal Markers of irony and sarcasm. *Humor – International Journal of Humor Research* 16(2). 243–260.
- Baken, Ronald J. & Robert F. Orlikoff. [1987] 2000 *Clinical Measurement of Speech and Voice*, 2nd edn. San Diego: Singular.
- Banse, Rainer & Klaus R. Scherer. 1996. Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology* 70(3). 614–636.
- Berg, Wolfgang. 1978. *Uneigentliches Sprechen: zur Pragmatik u. Semantik von Metapher, Metonymie, Ironie, Litotes und rhetorischer Frage*. Tübingen: Narr.
- Boersma, Paul & David Weenink. 2012. Praat. Doing phonetics by computer (version 5.3.04; 3 March 2014). <http://www.praat.org> (accessed 5 July 2018).
- Braun, Angelika & Christa M. Heilmann. 2012. *SynchronEmotion* (Hallesche Schriften zur Sprechwissenschaft und Phonetik 41). Frankfurt am Main: Peter Lang.
- Braun, Angelika & Annabelle Rosin. 2015. On the speaker-specificity of hesitation markers. *Proceedings XVIII ICPhS, Glasgow*. 10–14. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0731.pdf> (accessed 5 July 2018).
- Bryant, Gregory A. & Jean E. Fox Tree. 2002. Recognizing Verbal Irony in Spontaneous Speech. *Metaphor and Symbol* 17(2). 99–117.
- Bryant, Gregory A. & Jean E. Fox Tree. 2005. Is there an ironic Tone of Voice? *Language and Speech* 48(3). 257–277.
- Cheang, Henry S. & Marc D. Pell. 2008. The sound of sarcasm. *Speech Communication* 50(5). 366–381.
- Cheang, Henry S. & Marc D. Pell. 2009. Acoustic markers of sarcasm in Cantonese and English. *The Journal of the Acoustical Society of America* 126(3). 1394–1405.
- Cheang, Henry S. & Marc D. Pell. 2011. Recognizing sarcasm without language: A cross-linguistic study of English and Cantonese. *Pragmatics & Cognition* 19. 203–223.
- Chen, Aaju & Lou Boves. 2018. What’s in a word: Sounding sarcastic in British English. *Journal of the International Phonetic Association* 48(1). 57–76.
- Creusere, Marlena 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. 29–45.

- Cutler, Anne. 1974. On saying what you mean without meaning what you say. In La Galy, Michael W., Robert A. Fox & Anthony Bruck (eds.), *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, 117–127. Chicago, IL: Chicago Linguistic Society.
- Fónagy, Ivan. 1971. The functions of vocal style. In Chatman, Seymour B. (ed.), *The functions of vocal style*. 159–176. London: Oxford University Press.
- Gibbs, Raymond & Jennifer O'Brien. 1991. Psychological aspects of irony understanding. *Journal of Pragmatics* 16. 523–530.
- Gibbs, Raymond W. & Herbert Colston. 2007. *Irony in Language and Thought: A Cognitive Science Reader*. New York: Erlbaum.
- Hansen, John H.L. & Sanjay A. Patil (2007). Speech under Stress: Analysis, Modeling and Recognition. In Christian Müller (ed.), *Speaker Classification I: Fundamentals, Features, and Methods*, 108–137. Springer: Berlin & Heidelberg.
- Haverkate, Henk (1990). A speech act analysis of irony. *Journal of Pragmatics* 14. 77–109.
- Hildebrandt, Bruno. 1961. Die arithmetische Bestimmung der durativen Funktion. Eine neue Methode der Lautdauerbewertung. *Zeitschrift für Sprachwissenschaft, Phonetik und Kommunikationsforschung* 14. 328–336.
- Kienast, Miriam 2002. *Phonetische Veränderungen in emotionaler Sprechweise*. Berlin: Shaker.
- Laval, Virginie & Alain Bert-Erboul. 2005. French-Speaking Children's Understanding of Sarcasm: The Role of Intonation and Context. *Journal of Speech, Language and Hearing Research* 48. 610–620.
- Lævenbruck, Hélène, Mohamed Ameer Ben Jannet, Mariapaola D'Imperio, Mathilde Spini, Maud Champagne-Lavau. 2013. Prosodic cues of sarcastic speech in French: slower, higher, wider. In Bimbot, Frédéric, Christophe Cerisara, Cécile Fougeron, Guillaume Gravier, Lori Lamel, François Pellegrino & Pascal Perrier (eds.), *Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech 2013)*, Lyon, France, 3537–3541. <https://hal.archives-ouvertes.fr/hal-00931864> (accessed 5 July 2018).
- Low, Ee Ling & Esther Grabe. 1995. Prosodic patterns in Singapore English. In Elenius, Kjell & Peter Branderud (eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences (ICPhS)*, Stockholm, Sweden, Vol. 3, 636–639.
- Nauke, Astrid & Angelika Braun. 2011. The production and perception of irony in short context-free utterances. In *Proceedings XVIIth International Congress of Phonetic Sciences (ICPhS)*, Hongkong, 1450–1453.
- Pexman, Penny M. & Kara M. Olineck. 2002. Does Sarcasm Always Sting? Investigating the Impact of Ironic Insults and Ironic Compliments. *Discourse Processes* 33(3). 199–217.
- Probst, Louise & Angelika Braun. 2016. Geflüsterte Angst und behauchte Trauer – Stimmqualität und Emotionen. Paper presented at P&P Munich, October 13–14.
- Rao, Rajiv. 2013. Prosodic Consequences of Sarcasm Versus Sincerity in Mexican Spanish. *Concentric: Studies in Linguistics* 39(2). 33–59.
- Rockwell, Patricia. 2000. Lower, Slower, Louder: Vocal Cues of Sarcasm. *Journal of Psycholinguistic Research* 29(5). 483–495.
- Schaffer, Rachel. 1981. Are there consistent vocal cues for irony? In C. S. Masek, R. A. Hendrick & M. F. Miller (eds.), *Papers from the Parasession on Language and Behavior*, 204–210. Chicago: Chicago University Press.
- Schaffer, Rachel. 1982. *Vocal cues for irony in English*. PhD dissertation Ohio State University.

- Scharrer, Lisa, Ursula Christmann & Maja Knoll. 2011. Voice Modulations in German Ironic Speech. *Language and Speech* 54(4). 435-465. DOI: 10.1177/0023830911402608.
- Schmiedel, Astrid. 2017. Phonetik ironischer Sprechweise. Produktion und Perzeption sarkastisch-ironischer und freundlich ironischer Äußerungen. Berlin: Frank & Timme.
- Shipp, Thomas & Harry Hollien. 1969. Perception of the aging male voice. *Journal of Speech and Hearing Research* 12. 703–710.
- Wilson, Deirdre. 2013. Irony comprehension: A developmental perspective. *Journal of Pragmatics* 59. 40–56.
- Winner, Ellen, Gail Windmueller, Elizabeth Rosenblatt, Laurie Bosco, Edella Best & Howard Gardner. 1987. Making Sense of Literal and Nonliteral Falsehood. *Metaphor and Symbol* 2(1). 13–32.