University of Hormozgan

Iranian Journal of Educational Research

Print ISSN: 1735 - 563X Online ISSN: 2980 - 874X

Homepage: http://ijer.hormozgan.ac.ir

Educational and Behavioral Research Center

Audiovisual Representation of Multimodal Cues During the Production of Irony and Sarcasm in Persian Sitcom

Raha Koochacki[1], Mohammad Hossein Sharafzadeh[2]✉, Ameneh Zare[3]

1. PHD Student PHD Student, Department of Linguistics, Marv.C., Islamic Azad University, Marvdasht, Iran
2. Department of Linguistics, Marv.C., Islamic Azad University, Marvdasht, Iran, mh.sharafzadeh@iau.ac.ir
3. Department of Linguistics, Marv.C., Islamic Azad University, Marvdasht, Iran

Article Info

Article type:

Research Article

Article history:

Received 23 May. 2025

Received in revised form 10 Aug. 2025

Accepted 16 Sep. 2025

Published online 01 Mar. 2026

Keywords:

Sarcasm, Production, Audiovisual prosody, Prosodic-gestural patterns

ABSTRACT

Objective: Sarcasm is a complex communicative phenomenon that relies on multimodal cues, including prosodic and gestural signals. Understanding how these cues interact contributes to the broader framework of audiovisual prosody. This study aimed to explore the interaction between prosodic and gestural cues in sarcastic utterances, examining how these multimodal features co-occur and differ from non-sarcastic speech.

Methods: A corpus of 33 sarcastic utterances produced by a professional ironist in a semi-summary genre television monologue was analyzed. Each sarcastic utterance was compared to a non-sarcastic utterance immediately preceding it. Prosodic and gestural cues were systematically labeled using PRAAT and ELAN software. Quantitative analyses were complemented by an in-depth qualitative examination of four sarcastic utterances to investigate the interplay of lexical-syntactic, prosodic, and gestural patterns.

Results: Quantitative findings indicated that sarcastic utterances contained a higher density of prosodic and gestural cues compared to the preceding non-sarcastic utterances. Moreover, prosodic and gestural cues were found to appear both in alignment and independently, often functioning as gestural codas. Qualitative analysis revealed a diverse range of relationships between lexical-syntactic structures, prosody, and gestures in the production of sarcasm.

Conclusions: The study highlights the multimodal nature of sarcasm, demonstrating that prosodic and gestural cues jointly contribute to its expression. These findings support the audiovisual prosody framework and suggest that irony comprehension relies on complex cue integration across modalities.

Cite this article: Koochacki, R. M., Sharafzadeh, M. H. & Zare, A. (2026). Audiovisual representation of multimodal cues during the production of irony and sarcasm in Persian sitcom. *Iranian Journal of Educational Research*, 5 (1), 1-32. . DOI: https://doi.org/10.22034/5.1.1

Publisher: University of Hormozgan.

DOI: https://doi.org/10.22034/5.1.1

Introduction

Indirect language is one of the most commonly used resources within the complex system of human communication. According to Bryant (2011, 291), there is a linguistic act in which the superficial verbal constituents are not a reflection of the message that the speaker wishes to convey, or, in other words, in which the ultimate meaning of the expression is not contained only in the propositional form of the statement, but is inferred from the interaction between this and other factors, such as the knowledge shared between the interlocutors or the way in which the statement has been uttered. It is indirect language because there is an incongruity between the literal content of the proposition and the implicit content, and we talk about "irony" and "sarcasm" because speaker does not intend to guide the listener towards the correct interpretation of the message. It is in that will of the "ironist" that his intention is perceived by the listener where we find the key to understanding the role that prosody and gestures play in the production and interpretation of ironic statements. On the other hand, according to Scharrer et al. (2011), in addition to the propositional content of the utterance, the factors involved in the interpretation of ironic and sarcastic statements are of a different nature and can be classified into two broad categories: the first category: the "shared knowledge" existing between the issuer and the receiver, which refers to the common knowledge about the situational context, about the world and the general beliefs of the speakers, and the second category consists of communicative clues or signs that indicate the presence of irony and sarcasm, which can be "segmental verbal" (e.g., the use of specific adjectives or adverbs or a specific syntactic position of the elements in that statement), "non-segmental verbal" (e.g., vocal modulations) or "non-verbal" (e.g., facial expressions and gestures). What is addressed here in this study is the role of these non-segmental verbal, or so to speak “prosodic or supra-segmental” components and non-verbal features, i.e. gestures and their interaction in the production of irony and sarcasm in a professional and systematic acted speech. In this way, we sought an appropriate and logical answer to these questions: 1) How are prosodic and gestural cues used by a professional ironist to indicate the presence of an ironic intention during and after his stand-up performance? 2) What is the rate of appearance of prosodic and gestural features in the production of irony and sarcasm by a professional ironist?

3) Are prosodic and gestural features temporally aligned during and after the production of irony and sarcasm? 4) Did the ironist in this experiment use specific and identical strategies to indicate irony and sarcasm?

2. The prosodic and gestural components in the previous study of irony and sarcasm

2-1. Prosody: ironic tone

Many are the studies that have described the prosodic variations that are observed when comparing the ironic speech with the neutral, which is why it is assumed that the speaker modulates its prosodic production in order to facilitate the listener of the interpretation of irony (for ex, Bryant, 2010, 2011; Rockwell, 2000).

The complexity of the phenomenon and the great diversity of ironic effects that occur in human communication greatly complicate the task of establishing a solid characterization of the ironic tone (as Bryant, 2010 and 2011 concludes), so most studies have focused on the description and analysis of prosody of a subtype of irony. The subtype that has deserved more attention has been the "critical irony" or sarcasm (e.g., Attardo et al, 2003, in French; hailed "positive image irony" (Ruiz Gurillo, 2008, 51) –that one in the intention is not to criticize, but flatter - (Nakassis and Snedeker, 2002, In English; Anolli and others, 2002, in Italian)). These restrictions, whether theoretical or methodological, have allowed obtaining results that point to the existence of some specific prosodic characteristics of some subtypes of irony, as well as observing the affinities and discrepancies existing between the different subtypes. In these studies they have analyzed the variations of prosodic elements such as height (Peaks, contours and global or local height –p. e.g., Focalizations– of f_0); the Intensity (global or local –p. e.g., focused words–) or the duration (global, of focused words, of specific segments, pauses or syllable).

In all of them there are significant variations between the ironic tone of voice and the non- ironic in any or in several of the acoustic parameters analyzed. However, while the slowdown in the production of ironic utterances - this is, the increase in

the overall duration of the ironic utterance or some of its segments - seems to be a characteristic phenomenon of the ironic speech that appears to continually and consistently reviewed transverse in all studies carried out (e.g. Anolli and others, 2002; Bryant, 2010; Laval and Bert-Erboul, 2005; Padilla, 2011). However, it seems that the results differ in the direction of the height and intensity modulations (See Scharrer et al., 2011 for a comprehensive review of discrepancies between height and intensity values from studies conducted). In short, studies conducted to date about the prosody of irony seem to confirm that speakers modulate the tone of voice when they emit an ironic statement and that it contrasts with non -ironic speech, but, as can be seen from the differences - and even contradictions - existing between studies, not in a unique and unequivocal way.

2-2. Gestures

Research on the use of gestures in combination with speech suggests that both discursive modalities, verbal and gestural arise from the same conceptual structure through an integrated process of statement construction (McNeill, 1992, 2005). Thus, from this perspective it is argued that speech and gestures form a unique and unified system, and that gestures not only co-social with speech, but are semantically and pragmatically co-constantly, showing the congruence of forms and regularities systematic regarding its position and synchrony, and jointly forming the "final product" that speakers conceive in the design or construction of their statements (Goldin-Meadow, 2003; Kendon, 2004; McNeill, 1992, 2005). According to these researchers, this joint conformation does not imply that the gesture is always reduced with the content of the discourse, but that on many occasions complement - not only by addition, but also by restriction - its meaning. From this prism, most of the gestures that occur together with speech would be acting as metadiscursive markers or punctualizers, reflecting the pragmatic function of a statement in speech or providing indications about its structure.

As for the studies that have addressed the gestural component in the production and perception of the ironic statements, the first thing that can be said is that they are scarce and to proceed less systematic than those dedicated to prosody. Even with this, the investigations carried out show that ironic speech is frequently accompanied by gestures and facial expressions such as head movements,

eyebrows, mouth and arms, as well as other nonverbal elements, such as laughter or gaze. It should also be noted that the approach to the study of the gestural components of ironic speech has been carried out mainly from two different, although related perspectives: that which addresses its study from the analysis of ironic speech (which includes the use of ironic expressions) (Attardo et al, 2003; Caucci and Kreuz, 2012; Tabacaru and Lemmens, 2014), And that which, in reverse, focuses on the analysis of the expression of irony (among whose communicative goals is irony) (Attardo et al, 2003; Bryant 2011, 2012; Haiman, 1998; Hancock, 2004; Kreuz, 1996; Williams et al, 2009).

In summary, and collecting the above in previous points, What we are looking for here in relation to gesture is the contribution that gesture, either alongside prosody or independently, can make to the correct interpretation of an ironic phrase, where we believe it is essential to specify the nature of that contribution, to relate it to prosody, and to place both components in a pragmatic model. The model that tries to account for the complexity of the phenomenon of irony without relegating prosody and gestures to the scope of the extra - or of the para -, or at least not before having detailed the type of information that both can code, both jointly and independently, and the specific function that both elements can perform. In the next point we will collect the frameworks of theorists who believe they fit more to this goal.

3-2. Theoretical framework of the study

As a manifestation of indirect language, the phenomenon of verbal irony has been addressed from very diverse perspectives, thus giving rise to multiple and varied typologies depending on the approach and the chosen classification criteria. From the field of psycholinguistics, for example, it has been suggested that verbal irony is used to reach social and communicatively complex goals (Kreuz and Roberts, 1995; Leggit and Gibbs, 2000). In other studies, whose focus is directed towards

the social functions of irony - as is the case of Tinge Hypothesis[1] (Dews and others, 1995) -, the phenomenon has been addressed according to the nuances that the use of ironic expressions - well with critical or flattering intention - it prints in the final interpretation of the message. On the other hand, from pragmatic theories of cognitive orientation (e.g., the theory of relevance - hereinafter referred to as TR[2]- (Sperber and Wilson, 1986/1995), or the Pretense Theory[3] (Clark and Gerrig, 1984)), it has been treated to explain the phenomenon of verbal irony attending to the processes of production, perception and cognitive processing of the ironic statements. The experimental data presented in this article will be discussed in the final section from the perspective of the TR, because there are several authors who have made in this frame –Vidal, 2011a for prosodic components; Wharton, 2009 for nonverbal elements).

1-3-2. Relevance Theory (TR)

In our opinion, the explanatory power of the TR derives from considering that not all linguistic elements contribute in the same way to the interpretation of a statement. Thus, the TR advocates the existence of different levels of representation in which units operate whose contribution is of different nature: conceptual units- which contain information about representations - and procedural units- which provide information on how to operate with these representations -

[1] According to the tinge hypothesis, then, the literal meaning of irony is activated initially, either before or alongside the ironic meaning, and is retained in order to dilute either the criticism or the compliment.

[2] According to TR, processing units operate at different levels: the level of "lower explanations," in which units – e.g. determiners or verb tenses – guide the addressee towards identifying the explicit content that the speaker wants to communicate; the level of "illocutionary explanations" (or upper explanations), in which units – e.g. intonation patterns, lexical evidentiality markers – account for the speaker's illocutionary expression or attitude; and the level of "implicatures," in which units – e.g. discourse markers – indicate how to connect the propositional content with other information in the context (see Escandell-Vidal, 2011a for more information).

[3] According to pretense theory, "ironists can pretend to use the words of any person or type of person they wish, just as long as they can get the intended audience to recognize the pretense" (Clark & Gerrig, 1984, 124).

(Wilson and Sperber, 1993, 2). It is in this second group of units – the procedural ones – that the contribution of some of the prosodic and gestural characteristics that accompany speech would be inscribed. In this sense, there are already several studies framed in the relevant perspective that have proposed that prosodic modulations encode procedural instructions that guide inferential processes through the reduction of the range of possible interpretations of an utterance. (House, 1990, 2006; Clark and Lyndsey, 1990; Fretheim, 2002; Wilson and Wharton, 2006; Escandell-Vidal, 1998, 2011a, 2011b; Prieto et al., 2013).

As for the gestures and facial expressions, the attention being less deserved, the studies published to date (e.g., Wharton, 2009; of De Brabanter, 2010; Forceville, 2014) coincide both in the importance they give to these elements (They can even constitute the only signal that manifests the intention of the performer), as in the need to make a clear distinction between the different types of gestures that we produce when communicating, then, in the same way that happens with the prosodic elements, its nature can range from the symbolic (e.g., universal) to the conventional (e.g., linguistic). In this sense, the typological refinement of this model allows us not only to characterize and associate with a certain functional level the contribution of those prosodic and gestural patterns that fit the procedural functions described, but also that of those no they necessarily encode specific processing instructions, but undoubtedly guide the interpretation of an utterance. In close connection with the above, we believe it is appropriate to point out that within the TR framework, numerous studies have also been dedicated to the phenomenon of humorous communication (e.g., Yus, 1997, 2003; Ruiz Gurillo & Alvarado Ortega, 2013). Given the close relationship between both phenomena - irony and sarcasm - the results obtained in the present study in the light of these works will be briefly discussed.

2-3-2. Audiovisual prosody perspective

On the other hand, from the perspective of "audiovisual prosody" it is stated that the prosodic characteristics of speech are, at the very least, complemented by

gestural marks (Krahmer and Swerts, 2004; Swerts and Krahmer, 2005). Thus, there are recent works in which it has even been observed that gestures provide more conclusive clues than intonation when it comes to interpreting the pragmatic content of an utterance (Borràs-Comes et al., 2011; Prieto et al., 2011). Other studies (Cvejic et al., 2010, 2012) have obtained results indicating that speakers are capable of making a kind of abstract prosodic representation based on the visual cues or marks they obtain from their interlocutors, a circumstance that allows them to correctly interpret an utterance despite inter- and intra-subject differences. Along these lines, some more recent experiments have focused on the study of acceptability, the degree of specificity and the interpretation resulting from the combination of gestural configurations and intonational patterns (Borràs-Comes and Prieto, 2011). It seems that, on the one hand– and as expected–, the different combinations lead to different interpretations and, on the other hand, not all combinations are acceptable. In addition, some of these gestural configurations are more general than others (and, therefore, more compatible with the different intonational configurations), while others are much more specific and, therefore, less combinable. In light of these results, it has become clear that the combination of certain gestural and intonational configurations leads to different interpretations, which are related to different prosodic categories, an obvious fact that, we believe, is a clear stimulus to approach the study of the phenomenon of irony from this perspective.

We are not aware of any detailed study to date on when and how prosody and gestures interact in ironic and sarcastic utterances. The present paper, which is purely exploratory in nature –both by design and by extension–, merely aims to open that door by presenting the results of a simple case study carried out on a corpus of thirty-three sarcastic utterances, which were subjected to (1) a quantitative analysis, which served to characterize the production of global prosodic-gestural markers in the corpus, and (2) a qualitative analysis of two of the utterances, in which we focused on observing the synchronicities existing between prosodic, gestural and lexical markers, as well as determining the function that they all played, jointly or independently.

3. Methodology

To carry out the case study that we present below, we initially elaborate a corpus of sixty three ironic utterances, whose selection, filtering and subsequent analysis was carried out according to the criteria we described below.

The sixty three ironic utterances[4] were selected and extracted by the researcher of a total of five episodes[5] that contain as many ironic monologues. The presenter of these monologues (belonging to the semi -summary genre[6] of the television monologue) is the popular ironist Mehran Modiri, and the contextual framework is that of a sitcom program as called “Dorehami”. The choice of this genre is due to the fact that we believe that the characteristics of this genre favor, on the one hand, the specific appearance of the statements of a sarcastic character - since sarcasm is precisely one of the goals of verbal irony (see Attardo et al., 2011, 2013; Ruiz-Gurillo, 2013) - and, on the other hand, the almost certain appearance of sarcastic and gestural signs - given the acted and dramatic component of this genre. We believe that all this is not detrimental to the objective of the study - because this is not to characterize ironic speech in spontaneous situations - but, by contrary, precisely because it is a genre that take place in a situation and a very context determined, the variables that could affect the data enjoy greater control. In addition, as Attardo and others (2003, 246-247), the data extracted from literary texts or other non-spontaneous texts can become "as revealing as the data obtained naturally". In order to confirm the prototype of the selected sarcastic statements,

[4] We have delimited the unit of utterance following a criterion of a discursive nature, that is, attending to pragmatic reasons, and not grammatical ones. Following Escandell-Vidal 2006, p. 28, we have avoided the identification of utterance-sentence, considering that «a unit of discourse cannot have more limits than those established by the speaker and his communicative intention, regardless of the degree of complexity of its formal realization». Thus, those sentences –or series of sentences– whose communicative intention–to ironize– remained constant throughout the communicative act have been considered «ironic utterances».

[5] The episodes have an average duration of 5’23’’, and are available for free on the website http://www.youtube.com.

[6] They are mostly scripted, although the monologue presenter may deviate from the script, which often happens in the particular case in question. In any case, we do not believe that this fact substantially conditions the observations made about irony, since this is not usually scripted, but rather forms a substantial part of the monologue genre itself.

we ensured that all of them were adjusted to the definition of an ironic statement proposed by Wilson and Sperber (1992, 59-60)[7]. The variable «Subtype of irony« was not contemplated in the selection of the statements of the corpus, Therefore, all those expressions that meet the criteria mentioned above are considered sarcastic, as well as the general observations made by Gibbs (2000: 13) from the field of psychology, namely, "Any form of sarcasm clearly reflects the idea of a speaker and creates some kind of conflict between expectations and reality." were considered for this study.

In the next step, in order to validate the prototype of selected sarcastic statements, we conducted a perceptual test on four informants - three women and one man, aged between 27 and 35 and with a high level of university education - which involved assessing the degree of sarcasm (on a Likert scale - from 1 indicating a very low level of understanding to 5 indicating a very high level -) that they perceived in each of the presented statements individually. As a result of this filter, finally those statements that obtained a score of 4.5 or higher were selected for analysis, which reduced the corpus to 33 sarcastic statements.

Subsequently, using the free software Audacity (Audacity Team, 2014), we recorded audio files, with which 33 audio files (in wav format and 16 bits) contained the sound of sarcastic speech and 10 seconds of sarcastic sound before it, to allow for a proper comparison between sarcastic and non-sarcastic speech. The recordings were then analyzed using Pratt software (Boersma and Weenink, 2008), which is designed for acoustic analysis of speech.

4. Data Analysis

4-1. Quantitative analysis thirty-three utterances

[7] From the relevant perspective, ironic statements are considered to be a variety of "indirect quotation." "Indirect quotations" - as opposed to "direct quotations" - are those in which a statement is not reproduced exactly, but only its meaning. Furthermore, for the indirect mention of a "proposition," a "meaning," or a "thought" to be considered ironic, it must be expressed through a clear attitude of disapproval or rejection of its content.

1-4-1. Prosody

The quantitative analysis of the prosodic components of the thirty-three sarcastic utterances and the thirty-three non-sarcastic utterances consisted of the extraction of four acoustic parameters related to the fundamental frequency (f_0), amplitude and time (following the proposal of Bryant, 2010). From the former, the following were extracted: (1) the mean $f_0$of the utterance (in Hz) and (2) the variability of f_0 (i.e. the mean of the deviations of the $f_0$values from the mean f_0 at each point of the intonational contour of each of the utterances) (in Hz); as for amplitude, the values of (3) the mean amplitude (in dB) were extracted; and finally, as for time, we calculated (4) the mean syllable duration (MSD), that is, the total time taken to pronounce the utterance divided by the number of syllables in that same utterance (in milliseconds), whose value is jointly calculated for the separation between words, as well as syllabification and the significant lengthening of segments, which are phenomena related to duration that are analyzed in the complete and exhaustive study on the ironic prosody of Spanish by Padilla 2011. The data from the four extracted acoustic parameters were subjected to four statistical t-tests in order to determine the independence of the means, with the independent variable being Type of utterance (sarcastic versus non-sarcastic) and the four acoustic parameters (Mean f_0, $f_0$variability, Mean amplitude and MSD –i.e. mean syllable duration–) being the dependent variables.

2-4-1. Gestures

Gesture cues were manually annotated by the author using the computer program ELAN (Lausberg and Sloetjes, 2009) following the Gesture and Facial Expression Coding Manual (Allwood et al., 2005 and Nonhebel et al., 2004). The labeled gestural components are all those described in the bibliography as possible cues for irony, such as head and mouth movements, eye opening/closing, and hand gestures or movements, as well as laughter and gaze aversion (e.g., Attardo, 2003, 2011; Bryant, 2011; Rockwell, 2000; Tabacaru and Lemmens, 2014; Williams et al., 2009). All gestures were labeled during utterance production as well as in those

moments immediately following utterance production in which we consider the produced gesture to be clearly integrated into the intended speech act. Given the observed difference between experimental conditions in terms of utterance length, the number of labeled gesture signs was divided by the syllables uttered in each of those utterances, so that the comparison between sarcastic and non-sarcastic sentences was not influenced by the segmental content of one being greater than the other. The percentage adjustment means an 11% reduction in the data obtained from sarcastic utterances.

4-2. Qualitative analysis of four of the thirty-three utterances

The second analysis was qualitative in nature and included a detailed description of the speaker's prosodic and gestural features during the production of ironic and sarcastic expressions and interaction of them with the lexical-syntactic component as well as their function. For this purpose, the analysis of prosodic phenomena explained in the previous section was supplemented by an applied phonological analysis of the intonational patterns of the ToBI system. Also, in this second analysis, the detailed description of gestural signs mentioned in the previous section was supplemented by the applied classification of gestures explained in McNeil, 1992, and the meaning of gestures was adapted to the specific needs of our work. This classification is based on the criteria of form (manual configuration and paths) and meaning (the perceived relationship of the gesture to the content and to the discursive structure The four categories that McNeill (1992) creates in terms of form are: 1) non-conventional gestures, which direct attention to a specific object (using the arms or head); 2) conventional gestures, which are symbols with a shared meaning within a community – for example, the OK gesture; 3) representational gestures (symbolic and metaphorical), which refer to objects, actions or relationships through the recreation of form or movement; and finally, 4) rhythmic gestures and movements (impulsive gestures) which, despite having no explicit meaning, are, in terms of prototype, a reflection of the speaker's production of discursive or narrative structures.

In the case of the temporal stages of gesture execution, this phase is divided into three well-defined phases: preparation, impact (i.e. execution, where the highest point of extension and intensity is called the peak), and withdrawal. In the case of rhythmic gestures, the moments of highest intensity - the peaks - usually coincide with prosodic signs such as f_0 peaks, indicating a close relationship between prosodic and gestural components. Observing these alignments constitutes one of the main objectives of this second analysis.

5. Results

5.1. Results of the quantitative analysis of the 33 utterances

In Table 1, the results of the mean and standard deviation of the four observed acoustic parameters can be seen, both for the 33 sarcastic utterances and for the 33 non-sarcastic utterances that were produced immediately before the sarcastic utterance. The data of the four parameters were analyzed in four statistical t-tests with the independent variable “utterance type” (sarcastic vs. non-sarcastic) and the four acoustic parameters (mean f_0, mean $f_0$variability, mean amplitude and MSD - i.e. mean syllable duration -) as dependent variables, in order to determine the independence of the means. This statistical analysis indicates that only the acoustic parameters “$f_0$variability” and “MSD” significantly differentiate both utterance types ($p<0.5$).

Table 1. Mean and standard deviation of the values of the four acoustic parameters collected from the 33 previous sarcastic and non-sarcastic utterances. Mean $f_0$and $f_0$variability values are shown in Hz, mean amplitude values in dB, and MSD (mean syllable duration) values in milliseconds.

	Sarcastic Utterances		Non-sarcastic Utterances	
Acoustic Parameters	Mean	SD mean	Mean	SD mean

(Hz)f_0	157.17	19.15	148.09	16.40
$f_0$variability (Hz)	41.63*	25.80*	26.77*	12.21*
Amplitude Mean (dB)	63.21	5.01	59.68	3.12
MSD (ms)	229.67*	41.65*	203.26*	22.75*

Note: The "*" sign indicates statistically significant values in t-tests ($p< 0.05$).

As for the results of the gestural markers, Table 2 presents the results for each utterance type of the mean number of markers produced per utterance. The length (measured in syllables) of the selected sarcastic utterances turned out to be 10% longer than that of the non-sarcastic ones, so the values presented have been corrected to adjust the time in ms of both and ensure the reliability of the comparison between the two types of utterance (see Table 2).

Therefore, we observed how sarcastic statements are generally produced with a greater number of gestural cues compared to non-sarcastic statements. The results of multiple chi-square statistical tests for nominal variables between the variable “type of utterance” (sarcastic vs. non-sarcastic) and each of the labeled gestural signals (absence vs. presence (one to three occurrences) vs. frequency (more than four occurrences)) showed significant differences between both utterance types in the following markers: mouth grimacing (0.7 in sarcastic statements vs. 0.3 in non-sarcastic statements), frowning (1.8 vs. 0.2) and arching the eyebrows (7.2 vs. 2.4), head tilting (4.3 vs. 0.2), half-closed eyes (2.8 vs. 0.3), and laughing/smiling (4.8 vs. 1.3). Although not reflected in Table 2, it is worth noting that a higher number of sarcastic statements were observed with gestural productions after expressing of the segmental verbal content of the statement, i.e., gestural codas, (in 71% of sarcastic statements and in 21% of non-sarcastic ones).

Table 2. Mean gestural markers appeared in thirty-three sarcastic utterances and in thirty-three non-sarcastic utterances.

	Sarcastic U.	Non-sarcastic U.
Gestural markers	Mean	Mean
Head – Nod	5.2*	2.8*
Head - Tilted	4.3*	0.2*

Head - Shaking	0.3	0.2
Eyebrows – Arched	7.2*	2.4*
Eyebrows–Furrowed	1.8*	0.2*
Mouth - Grimace (stretched lips)	0.7*	0.3*
Eyes - Half closed	2.8*	0.3*
Smile/laugh	4,8*	1.3*
Gaze - Averted	3.5	0.8
Hands - Metaphorical gestures	3.6	0.5
Hands-Shaking (Beat gestures)	5.5*	3.7*
Shoulders - Shrugging	6.9*	0.9*

Note. The '*' sign indicates statistically significant values in the chi-square tests ($p< 0.05$).

In summary, what we observed in the results of the quantitative analysis was that sarcastic utterances are produced with a slightly higher mean $f_0$and mean amplitude—although not significantly—in sarcastic speech than in non-sarcastic speech, as well as with a significant difference in $f_0$and mean syllable duration (MSD) between the two conditions. We also observed how gestural cues generally appear in sarcastic utterances in a higher percentage than in non-sarcastic ones, and how both conditions show statistically significant differences in the frequency of occurrence of nine of the twelve markers.

5.2. Qualitative analysis of two sarcastic utterances

In the following section, as an example of qualitative analysis, we examine in detail the relationship between segmental verbal (e.g., lexical-syntactic), non-segmental verbal (e.g., prosodic), and non-verbal (e.g., gestural) cues by analyzing two of the 33 sarcastic utterances and two of 33 immediately preceding non-sarcastic in the form of one Proposition, titled "The Service Episode,". We also deal with the different functions that prosodic and gestural components seem to perform. Here are two versions of this proposition: one in Persian in form of Persian transcription, our target for all acoustic and gestural analyses and the other is an English version of it for a better understanding of the meaning of the content by non-Persian-speaking audiences.

Episode title: servicing

Phonetic transcription:

Non-sarcastic utterance immediately preceding: “Indʒür ädæmä mäšin hæm kə mixän toulid konæn, čär tä lästikə, **yə büq**!do tä äynə bæghæl bä yə kæmærbænd!”

Sarcastic utterance: “bæqi yæš äpšən mæhsüb mišə!”

Non-sarcastic utterance immediately preceding: “agə bəxäyn bäyæd püləšo bədın.”

Sarcastic utterance: “væli bə qeımæt ə monäsəb, ämäde yə tæhvil, dær ræng häye motənævé.”

English version:

Non-sarcastic utterance immediately preceding: “When the people like this want to produce car, it has four tires, one horn! Two side mirrors, and one seat belt!”

Sarcastic utterance: “The rest of it is considered an option!”

Non-sarcastic utterance immediately preceding: “If you want it, you have to pay for it.”

Sarcastic utterance: “But at a fair price, ready for delivery, in various colors!”

Analysis of the acoustic parameters of these two sarcastic utterances compared to the two baseline utterances (i.e., the non-sarcastic utterances immediately preceding it) revealed an increase in the mean f_0 of the first sarcastic utterance relative to the non-sarcastic utterance immediately preceding it (216.89 Hz vs. 177.41 Hz) and a decrease in the mean f_0 of the second sarcastic utterance relative

to the non-sarcastic utterance immediately preceding it (133.86 Hz vs. 218.68 Hz), but greater $f_0$variability in both sarcastic utterances (47.31 for the first sarcastic utterance vs. 31.43 and 41.34 for the second utterance vs. 38.42). The mean amplitude is also slightly higher in both sarcastic utterances (71.43 dB vs. 68.65 dB; and 70.63 vs. 67.12), and the mean syllable duration differs significantly between the two utterance types (200.6 ms vs. 157.6 ms and 279.5 ms vs. 177.1 ms). At the local level, several sounds are pronounced louder: the consonant [n] in the words "mäšin" (car), "mixän" (they want) and "toulid konæn" (they produce); the consonant [q] in the word "büq" (horn); the consonant [b] in the words "bæqi yæš" (the rest of it) and "bæghæl" (side), the vowel [ä] in " äpšən " (option); the consonant [p] in the word "püləšo" (pay) and the consonant [t] in the word "tæhvil" (delivery) (see Figures 1& 2).

As for the pitch patterns, it is worth noting that the proposition consists of five independent intonational phrases, which correspond to five statements of this proposition and most notable pitch accents fall on the stressed syllable of the words “büq”, “äpšən” and “püləšo”.

We see how the pitch accent used in the word "äpšən" is the typical pre-nuclear accent of a broad-focus declarative sentence (L* +> H L ٪) (see Figure 1), although when it occurs at such a high frequency, the perception of this neutral phonological meaning becomes somewhat ambiguous and a clear emphatic load is perceived in the displacement of this pitch peak. It is worth noting that in the second statement “bæqi yæš äpšən mæhsüb mišə!” (the rest of it is considered option!), three of the five f_0 peaks occurred at a higher frequency each time, which is quite unusual in a declarative sentence, unless, as in this case, it is produced with emphasis. Furthermore, the speaker prepares the audience for a sarcastic utterance by emphasizing the word "büq" in the sentence before the first sarcastic statement, and this is clearly evident in both non-sarcastic sentences in this proposition. As you can see, this change in tone in the utterance of the word "püləšo" in the second non-sarcastic statement is also a prelude to preparing the audience for another sarcastic statement, which the ironist achieves with his own proficiency.

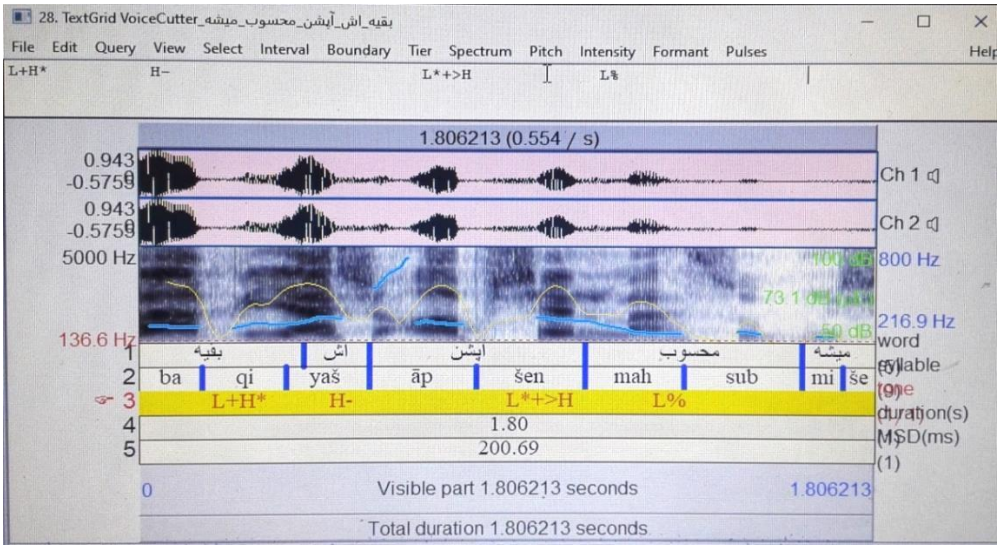

Figure 1. The output of waveform, spectrograph, and intonational contours from the sarcastic utterance: "bæqi yæš <u>äpšən</u> mæhsüb mišə" (The rest of it is considered an option!) in the PRAAT software. Oscillograms (top band), spectrograms (middle band), intonational contours (dark black lines in the top band), intensity curves (thin yellow lines in the middle band), fundamental frequency or pitch (thin blue lines in the middle band), utterance syllable fragmentation with phonetic transcription, syllable duration (in seconds), and mean syllable duration (MSD) of utterance (bottom).

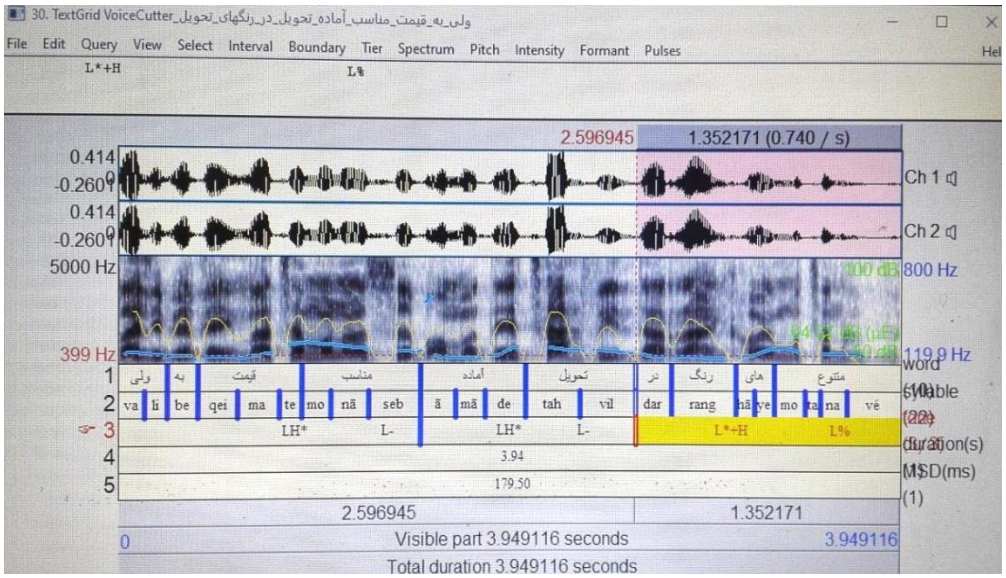

Figure 2. The output of waveform, spectrograph, and intonational contours from the sarcastic utterance: "væli bə qeımæt ə monäsəb, ämäde yə tæhvil, dær ræng häye motənævé" (But at a fair price, ready for delivery, in various colors!) in the PRAAT software. Oscillograms (top band), spectrograms (middle band), intonational contours (dark black lines in the top band), intensity curves (thin yellow lines in the middle band), fundamental frequency or pitch (thin blue lines in the middle band), utterance syllable fragmentation with phonetic transcription, syllable duration (in seconds), and mean syllable duration (MSD) of utterance (bottom).

Regarding gestures, we first of all observed how the speaker, most of the time, tilts his head slightly to the left during the utterance of the proposition and also shows a slight frown in the eyebrows, which is intensified at certain points, but will be present throughout the proposition. Hence, we were able to highlight the appearance of two different types of gestural signs: one of a rhythmic nature and the other of a conventional nature - which appeared in combination with a rhythmic gesture.

It also seemed that the peaks of rhythmic gestures coincided with the peaks of f_0 , and some of them were preceded by a portion of the speech slowdown. Thus, in the first sarcastic utterance, we observed alignments between the peaks of f_0 occurring in the stressed syllables of "bæqi yæš" and "äpšən" (produced with the emphatic pitch accents L+H* and L*+>H, see Figure 3) and the peaks of the following gestural signs:

Shaking the left hand to the left (during the production of the word "bæqi yæš"), raising the eyebrows, furrowing the eyebrows (during the production of the word" äpšən"), and at the end of the utterance as a gestural coda, a smile, (see Figure 3). In the second part of the sarcastic utterance (i.e., the utterance "væli bə qeımæt ə monäsəb, ämäde yə tæhvil, dær ræng häye motənævé " (But at a fair price, ready to be delivered, in various colors!), the rhythmic gestures appeared were: nodding the head forward in all three occurrences of the words "qeımæt ə monäsəb", "ämäde yə tæhvil", "dær ræng häye motənævé", raised eyebrows, only in the production of the word " monäsəb," and half-closed eyes, in the production of the word " ämäde yə tæhvil" and nodding the head left and right with opening and closing the eyes, in pronouncing the word "ræng häye motənævé" (see Figure 4).

In short, the overall contrasts between the two types of utterances seemed to be that both the prosodic features—for example, the $f_0$variability and MSD were significantly higher in sarcastic speech—and the gestural cues—such as the raised eyebrows and head tilt that were present most of the time during utterance production—were due to the speaker's overall distance from the literal content of the utterance he was pronouncing: the speaker's prosodic and gestural behavior were not the same.

Regarding specific phenomena, we also observed in this proposition how the peaks of rhythmic gestures (furrowed eyebrows, half-closed eyes, and shaking head and hand) were clearly aligned with $f_0$peaks, and most often in the parts where speech slowed down. In this proposition, in addition to observing the emphatic role of prosodic and gestural components in appearance of the words "äpšən", "qeɪmæt ə monäsəb", "ämäde yə tæhvil", and "ræng häye motənævé", we observed a different interaction between prsodic/gestural cues and the verbal- segmental content.

Thus, in the first part of the above proposition, we observed a gradual increase in three peaks of $f_0$in the production of the sarcastic utterance "bæqi yæš äpšən mæhsüb mišə!" (see Figure 3). Interestingly, the above utterance is quite unusual in the production of a declarative sentence and can only be explained if we look at the semantic content of the proposition and the communicative intention of the speaker. But in the second part of the proposition, the peaks of rhythmic gestures (i.e., nodding the head back and forth, raising the eyebrows, half-closing the eyes, and turning the head left and right) were clearly aligned with the F0 peaks, and most of the time in the parts where speech slowed down. Therefore, in the second sarcastic utterance, we observed alignments between the $f_0$peaks located in the stressed syllables of the words “monäsəb”, “tæhvil” and “ræng hä” (with the pitch accents L*H, LH* and L*+H) and the peaks of the gestural signs (see Figure 4).

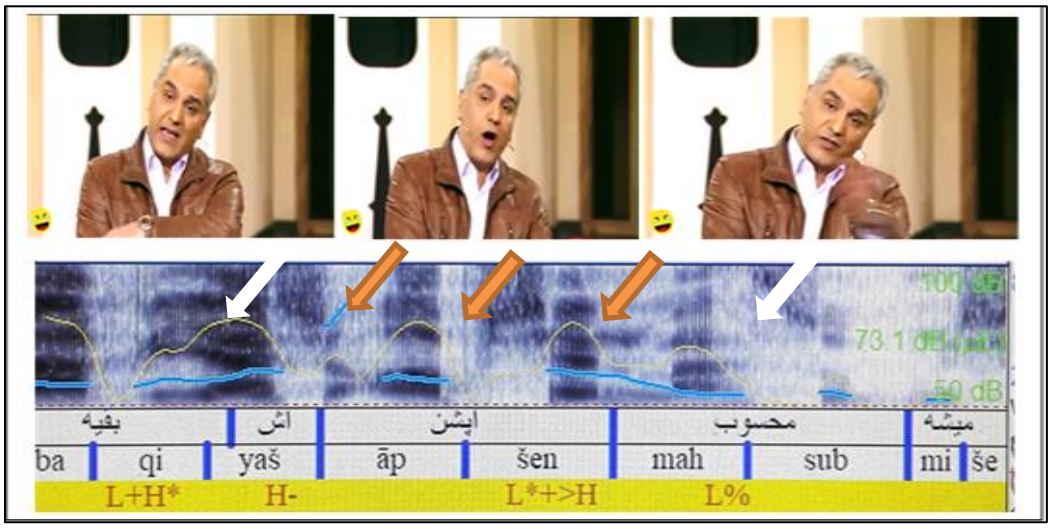

Figure 3. Image of the gestural sign produced during the production of the sarcastic utterance "bæqi yæš äpšən mæhsüb mišə!" accompanied by a rhythmic, left-striking arm gesture, the maximum extension of which corresponds to the point of greatest intensity of the eyebrows, half-closed eyes, and leftward head turn. All of which appeared in line with the F0 peak located in the stressed syllable of the above sarcastic words.

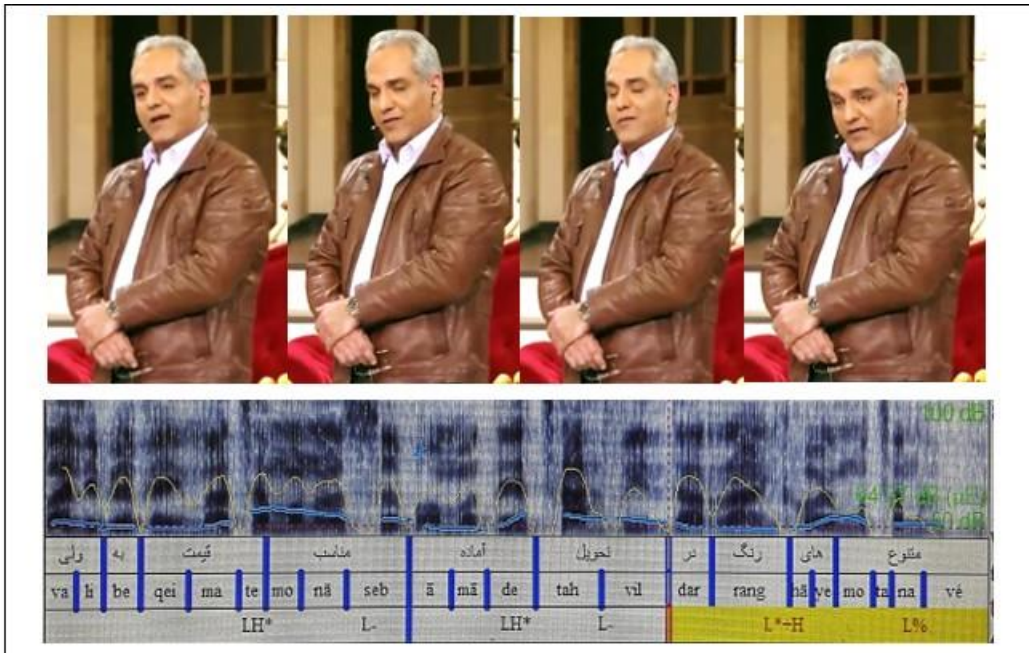

Figure 4. A representation of the gestural signal produced during the production of the sarcastic utterance "væli bə qeɪmæt ə monäsəb, ämäde yə tæhvil, dær ræng häye motənæv̲é", accompanied by a raised eyebrow followed by a forward nod of the head with each utterance of the sarcastic words and closing of the eyes, and finally a leftward tilt of the head with opening and closing of the eyes. All of them appear in line with the F0 peak located in the stressed syllable of the sarcastic words above.

6. Discussion and Conclusion

In the following section, we will first discuss the study findings in relation to the previous literature and then demonstrate how they contribute to the existing body of research in this field while answering the research questions.

Previous research on prosodic and gestural features of sarcasm has indicated that speakers convey prosodic and gestural components in their sarcastic speech. One of the issues that widely discussed in this literature has been whether there is a consistent tone of voice (or sarcastic gestural pattern) that we can identify in sarcastic speech.

Previous research in the scope of prosodic components has reported that sarcastic utterances are produced by acoustic modulations in pitch (with higher or lower mean fundamental frequency and fundamental frequency variability), intensity changes (with higher intensity values), and a diverse set of durational features (e.g., slower speech rate, more and longer pauses) (e.g., Gibbs, 2000; Nakassis & Snedeker, 2002; Loevenbruck et al., 2013; Anolli et al., 2002; Attardo et al., 2003, 2011; Laval & Bert-Erboul, 2005; Cheang & Pell, 2009; Bryant & Fox-Tree, 2002, 2005; Bryant, 2010; Scharrer et al., 2011; Padilla, 2004, 2011; Gonzalez-Fuente et al., 2016). Pitch and intensity cues have yielded different results across studies and languages, but the only feature of sarcasm that has been consistent across languages being a slower speech rate (or, in other words, a longer speech duration). Furthermore, some features of intonation have been reported to be associated with the speaker's sarcastic intent, such as the rising-final morphological patterns in Spanish (e.g., Padilla, 2004, 2009) or specific nuclear configurations in tone of voice for French (e.g., Gonzalez-Fuente et al., 2016).

Regarding visual cues, previous research has shown that speakers use a wide range of gestural cues when expressing sarcasm, for example, raising eyebrows, head movements, stretched lips, as well as smiling, laughing, and staring (Attardo et al., 2003, 2011; Bryant, 2011, 2012; Haiman, 1998; Hancock, 2004; Kreuz, 1996; Caucci & Kreuz, 2012; Gibbs, 2000; Williams et al., 2011; Padilla, 2004). In general, these gestural cues have been reported to convey information about the speaker's feelings and attitudes and to engage in a wide range of different socio-communicative functions, such as reinforcing a shared positive emotional experience (Smosky & Bachorowski, 2003), strengthening friendship bonds (Alvarado & Padilla, 2010), or criticizing something or someone (Sperber & Wilson, 1986/1995).

Considering the prosodic components, the results of the present study revealed that sarcastic utterances were marked by significantly higher fundamental frequency variability and slower speech rates than non-sarcastic utterances. These findings are consistent with previous studies reporting the use of intonational patterns to indicate sarcasm in different languages (e.g., Attardo, 2001, for English; Padilla, 2004, 2009, 2011, for Spanish; Gonzalez-Fuente et al., 2016, for French).

However, the use of specific tone-core configurations in sarcasm remains largely unknown. In this regard, recent research on the pragmatic meanings of intonation has indicated that specific tone-core configurations are closely related to the discourse functions of insubordinate and subordinate clauses (Elvira-García, 2016; Elvira-García, Roseano, & Fernández-Planas, 2017), strongly suggesting that much more research is needed to acheive a quite comprehending of the role of specific tone-core configurations in indicating a variety of linguistic meanings. Regarding gestural cues, the present study showed that speaker produced sarcastic utterances with a higher rate of gestural cues (e.g., eyebrow raising, head movements, smiling, laughing, and gaze shifts) compared to non-sarcastic utterances.

Taken together, the results of present experiment indicated that (a) ironist can indicate their sarcastic intent by combining a variety of prosodic and gestural elements, and (b) there is no single, unique way to indicate sarcastic intent through prosodic and gestural components, which leads us to conclude that we cannot identify a specific "sarcastic tone of voice" or a "sarcastic gestural pattern" that is specific to the display of sarcasm. Previous research has shown that the different and even contradictory results in studies examining "ironic tone of voice" can be explained by differences in methodological design in the irony subtype under analysis, in the specific implementation of the ironic language, and in the specific intonational phonology of each language (Bryant, 2011; Loevenbruck et al., 2013). However, the researcher in the other experiment revealed that there was no significant relationship between the "irony subtype" and the multimodal cues conveyed by the speakers, meaning that we did not find a specific or unique tone of voice for ironic, "sarcastic," or "hyperbolic" utterances.

Therefore our results suggest that even when fully controlling for "subtypes of irony" at the time of their analysis, searching for a consistent sarcastic tone of voice or a sarcastic gestural pattern leads to inconsistent results. Importantly, we believe that this is because the intentions and emotions that can be conveyed through a sarcastic remark range from highly positive to highly negative (Wilson, 2013; Yus, 2016), and the gestural and prosodic signals that convey these specific

emotions and attitudes are highly diverse and overlap with a wide range of communicative purposes. We argue that prosodic patterns (as well as gestural patterns) in different languages are particularly well suited to conveying intention and can therefore be highly diverse. In the case of gestures, speakers typically use a variety of smiles and laughs to convey positive intentions, while they use head shaking or frowning to convey negative intentions (see wharton, 2009). Furthermore, it is worth noting that the wide range of visual cues that can appear during communication can be multi-meaning (Poggi, 2007) and that the same eyebrow configuration combined with a different head movement can lead to different or even conflicting interpretations. It is therefore important to conduct further studies that consider the complexity of these gestural and prosodic patterns in communication.

What is evident from the qualitative analysis of two sarcastic utterances is that the speaker's intention in producing these sarcastic expressions is to emphasize the dishonesty of the statements by producing irony and sarcasm. Thus, first he states the first sentence of proposition (When the people like this want to produce car, it has four tires, one horn! Two side mirrors, and one seat belt!) and then, after the sarcastic use of the utterance "the rest of it is considered an option" (which is easily marked as a prosody and gesture, as explained earlier), and through the unusual use of a gradual increase in F0 peaks (synchronizing each of them with different rhythmic and marked gestures), as well as a decrease in the overall speed of speech, the speaker creates a discursive tension that reaches its peak with the use of the words "fair price, ready to deliver and various colors" in the second part of the proposition, which indicates the absurdity and dishonesty that the speaker wants to point out.

Furthermore, in the second sarcastic statement of this proposition, we observed a conventional gesture (nodding head) with a clear semantic content (agreement or approval) whose role is to exaggerate the positive evaluation that the speaker makes of the exaggerated expression, namely, "but at a fair price, ready to delivery, in various colors." Hence, it highlights the contrast between what is said, "but at a fair price, ready to be delivery, in various colors," and what is real—which they are

not—and thus encourages the sarcastic interpretation of the expression on the other side and creates the intended ironic effect.

Another important finding of this study concerns the temporal alignment between sarcastic and gestural cues. The results of both studies showed that sarcastic and gestural cues can be either temporally aligned or non-aligned with speech. In this regard, the main findings were: (a) that pitch stresses associated with emphatic nuclear tonal arrangements (e.g. L+H* L%, L+H*, L!H % and L*!H %) are often temporally associated with gestures and some facial movements, especially eyebrow and head movements, and (b) that some gestural cues appear independently of gestural cues, especially in what we call gestural codas or gestural patterns that appear in post-speech position, e.g. when gestural and lexical information are absent. Therefore, as a research suggestion, researchers can conduct a more extensive study in future studies on the presence of gestural cues in the time after ironic utterances and evaluate the extent to which this category of gestural cues (in other words, gestural codas) influence the perception of ironic utterances or other complex human communication in general.

Overall, the strong presence of gestural and prosodic cues (and especially gestural codas) in sarcastic speech found in our study highlights the importance of multimodal cues in sarcastic communication. These findings are consistent with proposals in relevance theory that gestural and prosodic cues are used by speakers to reduce the receiver's processing effort, ostensibly until the speaker is confident that the sarcastic comprehension process has been completed (e.g., House, 1990, 2006; Clark & Lindsey, 1990; Fretheim, 2002; Wilson & Wharton, 2006; Escandell-Vidal, 1998, 2011a, 2011b; & Wharton, 2009). Furthermore, the existence of gestural codas helps to clarify and explain the concept of “ironic speech” and emphasizes the claim made by functional narratives that an ironic speech act does not end until all relevant information has been expressed; in other words, a discourse unit, regardless of the degree of complexity of its formal realization, has no more constraints than those set by the speaker and his communicative intention” (Escandell-Vidal, 2006: 28).

7. References

Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C., & Paggio, P. (2005). The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing Phenomena. *Language Resources and Evaluation, 41*(3/4), 273-287.

Alvarado Ortega, M.B., & Padilla, X. (2010): Being polite through irony. In Dale April & Lidia Rodríguez (Eds.): *Dialogue in Spanish. Studies in functions and contexts. Dialogue Studies*, (pp. 55-68). Amsterdam: John Benjamins.

Anolli, L., Ciceri, R., & Infantino, M.G. (2002). "From 'blame by praise' to 'praise by blame': Analysis of vocal patterns in ironic communication". *International Journal of Psychology*, *37*, 266-276. 258

Attardo, S., Wagner, M. M., & Urios-Aparisi, E. (2013). "Prosody and Humour". In S. Attardo, M. M. Wagner, & E. Urios-Aparisi (Eds.), *Prosody and Humor* (pp. 189-201). Amsterdam: John Benjamins.

Attardo, S., Pickering L., & Baker A. (2011). Prosodic and multimodal markers of humor in conversation. *Pragmatics & Cognition*, *19*(2), 224-247.

Attardo, S., Eisterhold, J., Hay, J., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. *International Journal of Humor Research*, *16*, 243-260.

Attardo, S. (2001) *Humorous Texts: A Semantic and Pragmatic Analysis*. Berlin: Mouton de Gruyter.Austin, J. L. (1962): How to do things with words, Oxford, Clarendon.

Audacity Team (2014). Audacity(R): Free Audio Editor and Recorder [Computer program]. Version 2.0.0 retrieved April 20th 2014 from http://audacity.sourceforge.net/

Boersma, P., & Weenink, D. (2008). PRAAT: doing phonetics by computer. Version 5.4.08. Computer program. Retrieved 2 April 2015 from http://www.praat.org/.

Borràs-Comes, J., & Prieto, P. (2011). 'Seeing tunes'. The role of visual gestures in tune interpretation. *Journal of Laboratory Phonology*, *2*(2), 335-380.

Bryant, G.A. (2012). Is verbal irony special? *Language and Linguistics Compass*, *6*(11), 673-685.

Bryant, G.A. (2011). Verbal irony in the wild. *Pragmatics and Cognition*, *19*(2), 291-309.

Bryant, G.A. (2010). Prosodic contrasts in ironic speech. *Discourse Processes*, *47*(7), 545-566.

Bryant, G.A., & Fox Tree, J.E. (2005). Is there an ironic tone of voice? *Language and Speech, 48*(3), 257-277.

Bryant, G.A., & Fox Tree, J.E. (2002). Recognizing verbal irony in spontaneous speech. *Metaphor and Symbol*, *17*(2), 99-117.Cartmill E.A., Ece Demir, Ö. & Goldin-Meadow, S. (2012): Studying gesture, research methods in child language: A practical guide, Oxford, Wiley-Blackwell.

Caucci, G. M., & Kreuz, R. J. (2012). Social and paralinguistic cues to sarcasm. *Humor: International Journal of Humor Research*, *25*, 1-22.

Cheang, H. S., & Pell, M. D. (2009). Acoustic markers of sarcasm in Cantonese and English. *Journal of the Acoustical Society of America*, *126*(3), 1394-1405.

Clark, H. H., & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of Experimental Psychology: General*, *113*(1), 121-126.

Clark, B., & Lyndsey, G., (1990). Intonation, grammar and utterance interpretation. In *UCL Working Papers in Linguistics* 2 (pp. 32-51).

Cvejic E., Kim J., & Davis, C. (2012). Recognizing prosody across modalities, face areas and speakers: Examining perceivers' sensitivity to variable realizations of visual prosody. *Cognition, 122*(3), 442-453.

Cvejic, E., Kim, J., & Davis, C. (2010). Abstracting visual prosody across speakers and face areas, In *Proceedings of AVSP 2010*, Hakone, Kanagawa, Japan.

De Brabanter, P. (2010). "Uttering sentences made up of words and gestures". In Soria, B. & E. Romero (Eds.), *Explicit Communication: Robyn Carston's Pragmatics* (pp. 199-216). Basingstoke: Palgrave MacMillan.

Dews, S. & Winner, E. (1995). "Muting the meaning: A social function of irony". *Metaphor and Symbolic Activity*, *10*, 3-19.

Elvira-García, W. (2016). *La prosodia de las construcciones insubordinadas conectivo-argumentativas en español* (PhD disseration). Universitat de Barcelona, Barcelona, Spain.

Elvira-García, W., Roseano, P., & Fernández Planas, A. M. (2017). Prosody as a cue for syntactic dependency. Evidence from dependent and independent clauses with subordination marks in Spanish. *Journal of Pragmatics, 109, 29-46.*

Escandell-Vidal, M.V., & Leonetti, M. (2014). Fronting and Irony in Spanish. In A. Dufter & A. Octavio de Toledo (Eds.*), Left Sentence Peripheries in Spanish: Diachronic, Variationist and*

Typological Perspectives 3 (pp. 309-342). Amsterdam: John Benjamins.

Escandell-Vidal, M.V. (2011a). Prosodia y pragmática. *Studies in Hispanic and Lusophone Linguistics 4* (1), 193-208.

Escandell-Vidal, M. V. (2011b). *Verum focus* y prosodia: cuando la duración (sí que) importa. *Oralia*, *14*, 181-202.

Escandell-Vidal, M.V. (2006). Barcelona: Ariel. Introducción a la pragmática

Escandell-Vidal, M.V. (1998). Intonation and procedural encoding: the case of Spanish interrogatives. In Rouchota, Villy, Jucker & Andreas (Eds.), *Current Issues in Relevance Theory* (pp. 169-203). Amsterdam: John Benjamins.

Forceville, C. (2014). Relevance Theory as model for analyzing visual and multimodal communication. In D. Machin (Ed.), *Visual Communication* (pp. 51-70). Berlin: Mouton de Gruyter.

Fretheim, T. (2002). Intonation as a constraint on inferential processing. In *Proceedings of Speech Prosody International Conference*. Aix-en-Provence, France.

Gibbs, R. W. (2000). Metarepresentations in staged communicative acts. In D. Sperber (Ed.), *Metarepresentations: A Multidisciplinary Perspective [Vancouver Studies in Cognitive Science 10]* (pp. 389-410). New York: Oxford University Press.

Goldin-Meadow, S. (2003). *Hearing Gesture: How our hands help us think*. Cambridge: Harvard University Press.

González-Fuente, S., Prieto, P., & Noveck, I. (2016). A fine-grained analysis of the acoustic cues involved in verbal irony recognition in French. In *Proceedings of the Speech Prosody 2016* (pp. 902-906). Boston, MA (USA).

Haiman, J. (1998). *Talk is cheap: Sarcasm, Alienation, and the Evolution of Language.* New York: Oxford University Press.

Hancock, J. T. (2004). Ironic use in face-to-face and computer-mediated conversation. *Journal of Language and Social Psychology*, 23, 447-463.

House, J. (2006). Constructing a context with intonation. *Journal of Pragmatics*, *38*, 1542-1558.

House, J. (1990). Intonation structures and pragmatic interpretation. In S. Ramsaran (Ed.),*Studies in the pronunciation of English* (pp. 38-57). London: Routledge.

Kendon, A. (2004). *Gesture: Visible Action as Utterance.* UK: Cambridge University Press.

Krahmer, E., & Swerts, M. (2004). More about brows. In Z. Ruttkay & C. Pelachaud (Eds.), *From brows to trust: Evaluating embodied conversational agents* (pp. 191-216). Dordrecht: Kluwer Academic Press.

Kreuz, R. J. (1996). The use of verbal irony: Cues and constraints. In J. S. Mio and A. N. Katz (Eds.), *Metaphor: Implications and Applications* (pp. 23-38). Mahwah, NJ: Lawrence Erlbaum.

Kreuz, R. J., & Roberts, R. M. (1995). Two cues for verbal irony: Hyperbole and the ironic tone of voice. *Journal of Experimental Psychology: General 118*, 372-386.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior research methods, Instruments & Computers, 41(3),* 841-849. Computer program. Retrieved 14 February 2015 from http://www.lat-mpi.eu/tools/elan/.

Laval, V., & Bert-Eboul, A. (2005). French-speaking children's understanding of sarcasm: The role of intonation and context. *Journal of Speech, Language, & Hearing Research*, *48*, 610-620.

Leggit, J. S., & Gibbs, R. W. (2000). Emotional reactions to verbal irony. *Discourse Processes, 29*, 1-24.

Loevenbruck, H., BenJannet, M., D'Imperio, M., Spini, M., & Champagne-Lavau, M. (2013). Prosodic cues of sarcasm speech in French: slower, higher, wider. In *Proceedings of Interspeech 2013*. Lyon, France.

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago: University of Chicago Press.

Nakassis, C., & Snedeker, J. (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.

Nonhebel, A., Crasborn, O., & van der Kooij, E. (2004). Sign language transcription conventions for the ECHO project. Available at http://www.let.kun.nl/sign-lang/echo/docs/transcr_conv.pdf. Retrieved October 25th, 2012.

Padilla, X. (2011). ¿Existen rasgos prosódicos objetivos en los enunciados irónicos? *Oralia, 14*, 203-224.

Padilla, X. (2009): Marcas acústico-melódicas: el tono irónico. In L. Ruiz Gurillo & X. Padilla, X. (Eds.) *Dime cómo ironizas y te dire quién eres* (pp. 371-390). Frankfurt: Peter Lang.

Padilla, X. (2004). El tono irónico. Estudio fonopragmático. *Español Actual*, *81*: 85-98.

Nonhebel, A., Crasborn, O. & van der Kooij, E. (2004): Sign language transcription conventions for the ECHO project. <http://sign-lang.ruhosting.nl/echo/docs/ECHO_transcr_conv.pdf>. Descarga: 25/10/2012.

Poggi, I. (2007). *Mind, hands, face and body. A goal and belief view of multimodal communication*. Berlin: Weidler.

Prieto, P., Borràs-Comes, J., Tubau, S., & Espinal, M.T. (2013). Prosody and gesture constrain the interpretation of double negation. *Lingua*, *131*, 136-150.

Rockwell, P. (2000). Lower, slower, louder: Vocal cues of sarcasm. *Journal of Psycholinguistic Research*, *29*, 483-495.

Ruiz Gurillo, L. (2013): Narrative strategies in Buenafuente's humorous monologues. In , L. Ruiz Gurillo & M. B. Alvarado Ortega (Eds.), *Irony and humor: From Pragmatics to Discourse* (pp. 107-140). Amsterdam: John Benjamins.

Ruiz Gurillo, L. (2008). Las metarrepresentaciones en el español hablado. *Spanish in Context, 5*(1), 40-63.

Ruiz Gurillo, L., & Alvarado Ortega, M. B. (Eds.) (2013). *Irony and Humor. From pragmatics to discourse*. Amsterdam: John
Benjamins.

Scharrer, L., Christmann, U., & Knoll, M. (2011). Voice Modulations in German Ironic Speech. *Language and Speech*, *54*(4), 435-465.

Smoski, M. J., & Bachorowski, J. (2003). Antiphonal laughter between friends and strangers. *Cognition & Emotion*, *17*, 327-340.

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and Cognition*. Oxford : Blackwell.

Swerts, M., & Krahmer, E. (2008). Facial expressions and prosodic prominence: Comparing modalities and facial areas. *Journal of Phonetics*, *36*(2), 219-238.

Swerts, M., & Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language*, *53*, 81-94.

Tabacaru, S., & Lemmens, M. (2014). Raised eyebrows as gestural triggers in humour: the case of sarcasm and hyper-understanding. *European Journal of Humour Research*, *2*, 11-31.

Wharton, T. (2009). *Pragmatics and Nonverbal Communication*. Cambridge: Cambridge University Press.

Williams, J. A., Burns, E. L., & Harmon, E. A. (2009). Insincere utterances and gaze: Eye contact during sarcastic statements. *Perceptual and Motor Skills*, *108*(2), 565-572.

Wilson, D. (2013). Irony comprehension: A developmental perspective. *Journal of Pragmatics*, *59*, 40-56.

Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics, 38,* 1557-1579.

Wilson, D., & Sperber, D. (1993). Linguistic form and relevance". *Lingua* 90, 1/2, vol. 2, pp.1-25.

Wilson, D. & Sperber, D. (1992). On verbal irony. *Lingua*, *87*, 53-76.

Yus, F. (2016). Propositional attitude, affective attitude and irony comprehension. *Pragmatics & Cognition*, *23*(1), 92-116.

Yus, F. (2003). Humour and the search for relevance. *Journal of Pragmatics*, *35*(9), 1295-1331.

Yus, F. (1997). La teoría de la relevancia y la estrategia humorística de la incongruencia-resolución. *Pragmalingüística*, *3-4*, 497-508.