PENSOFT.

Languages and Modalities

Exploring the role of the body in communicating ironic stance

Clarissa de Vries[1], Bert Oben[1], Geert Brône[1]

1 *Department of Linguistics, Leuven University, Leuven, Belgium*

Corresponding author: Clarissa de Vries (clarissa.devries@kuleuven.be)

Abstract

Performing and understanding conversational irony requires a complex management of multiple viewpoints. To communicate and negotiate these intricate viewpoint shifts, speakers (and addressees) often use nonverbal means (e.g. gaze shifts, shrugs, shifts in body orientation, hand gestures, etc.) next to verbal viewpoint strategies. In the present paper we zoom in on the perspective of the speaker and try to describe and quantify bodily behavior in ironic utterances compared to non-ironic ones. To this end, we use data from a video-corpus of three-party interactions with participants wearing mobile eye-tracking devices that allow for precise eye gaze data. Our results show that speakers display more of the multimodal resources under scrutiny in ironic cases compared to non-ironic cases. More specifically, the involvement of bodily resources is mainly manifested in the use of laughter, head movements and body repositionings. We further show how those resources cluster into certain multimodal packages, and how the exact timing of the bodily behavior is relevant (i.e. the gaze behavior at the end of an ironic segment differs most notably from the end of a non-ironic one). Next to a quantitative analysis of the resources used in ironic talk in interaction, we also illustrate our findings with qualitative descriptions of relevant examples.

Key Words

1. Introduction

It is a broadly accepted view that our language system, and more generally our cognitive system, allows us to organize, communicate and access information from different viewpoints, which can be temporal, spatial, epistemic or other. The strongly 'viewpointed' nature of language (Dancygier & Sweetser, 2012) can be attributed to the foundations of embodied cognition, which hold that abstract cognition is rooted in and shaped by aspects of the human body. One such shaping factor is the fact that *by default* we perceive and process the world around us from a particular angle. What makes the human cognitive system unique, however, is its capacity to switch between perspectives or access multiple viewpoints at the same time (Dancygier & Vandelanotte, 2016; Duijn & Verhagen, 2019; Sweetser, 2012; Tomasello, 2008; van Krieken et al., 2019; Vandelanotte, 2019; Zeman, 2019). Being

able to (re)view a scene from different angles, including that of others (Theory of Mind), the viewpoint of oneself in the past, future or some counterfactual space, or that of a character in a fictional world, allows for complex cognitive representations to emerge and high-level communication on such representations to take place. As a matter of fact, this capacity for managing multiple viewpoints is built into the grammatical system, including viewpoint markers and multiple perspective constructions (see Dancygier, 2017 for an overview).

Research in a variety of disciplines, including cognitive science, linguistics, literary analysis and philosophy, has focused on the ways in which viewpoint permeates all products of human cognition, from mundane talk-in-interaction over intricate narratives to physical works of art (Dancygier et al., 2016; Dancygier & Sweetser, 2012; McIntyre, 2006; Tobin, 2018; Vandelanotte, 2017, and many others). In recent work, particular attention

has been paid to the question of how the handling of or shifts between multiple viewpoints may be marked multimodally, in both spoken and signed language. In other words, what semiotic resources do language users draw on to distinguish between and navigate through different layers of action? Among the pioneering studies on this topic are Sidnell (2006) and Thompson & Suzuki (2014), which have shown that in storytelling-in-interaction, speakers tend to indicate a shift from the original event of the telling (the narrator perspective) to the habitat of the reenacted event (the character perspective in the story) by means of eye gaze. More specifically, whereas speakers-as-narrators tend to direct their gaze (at least partially) towards their addressees, they systematically avert their gaze from the audience during reenacted sequences, in which they adopt the character perspective. Along the same lines, Parrill (2012) and Sweetser & Stec (2016) provide evidence that gaze, in combination with head movements and gestures, may indicate viewpoint shifts that are crucial to a correct parsing of a telling in interactionally relevant units. What these studies show, is that different semiotic resources may be linked to different layers of action (e.g. the gaze and head orientation may be associated with a character viewpoint, whereas the torso may be associated with a narrator oriented towards his/her audience), thus realizing a form of body partitioning for representing multiple viewpoints.

The above-mentioned studies all deal with anecdotes and longer narrative sequences in talk-in-interaction. In his seminal work *Using Language*, Clark (1996) also identifies phenomena in which this layering takes place within a single communicative act, referred to as *staged communicative acts* (ibid: 368). These include phenomena such as teasing, irony, sarcasm and jocularity, in which speakers briefly set up locally improvised scenes comparable to larger layered structures, but restricted in terms of their 'lifespan' in the ongoing interaction. Let's take, by way of illustration, the often-cited example in (1), discussed by Clark (ibid.: 353).

(1) Ken: and I'm cheap, - - -
     Margaret: I've always felt that about you,
     Ken: oh shut up,
                     (- - laughs) fifteen bob a lesson at home, -

In this specific sequence, a couple (Ken and Margaret) is engaged in a casual conversation on the husband's work as a private teacher. When he makes a statement on the fee he charges as a tutor (*and I'm cheap*), his wife reacts with what at the surface level may seem as a confirmation (*I've always felt that about you*), using the anaphoric pronoun *that* to explicitly link her utterance to the preceding one. However, it is immediately clear that she construes a different meaning of the adjective *cheap* for ironic-playful purposes. When embedded in the phrasal construction *to feel something about someone*, this meaning shifts to an extended reading 'of low moral value' rather than the literal 'inexpensive'. Underlying such an apparently simple case of wordplay, in fact, is an intricate play on viewpoints. In formulating her teasing reply, Margaret does not intend to seriously categorize her husband in any negative way. Rather, she sets up a pretense reading in which she reacts as if Ken had used *cheap* self-disparagingly in the extended sense, and she responds affirmatively. In other words, the ironic effect results from a tension between two layers of action and corresponding viewpoints: at the level of the actual communicative interaction, Margaret pretends that she, at a second level, seriously claims that Ken is metaphorically cheap. Thus, in the brief improvised scene, Margaret adopts the role of an implied participant, with different views from the perspective of which the scene is (re)viewed.

Many studies that zoom in on humor and irony in (face-to-face) interaction adopt some version of a pretense-based view on the phenomenon, like the one presented by Clark (and originally developed in Clark & Gerrig (1984). For instance, this view is explicitly or implicitly integrated in the cognitive-linguistic accounts such as those presented by Barnden (2017), Brône (2008, 2021); Brône & Oben (to appear), Coulson (2005), Geeraerts (2021), Kihara (2005), Tobin & Israel (2012) and many others. Crucial to any pretense-view on irony is that all parties involved, be it a writer and their readers or a speaker and their addressees, manage to see through the pretense, in the sense that they manage to access the multiple viewpoints involved and identify the relevant contrasts between those viewpoints. This is apparent in example (1), where the husband laughingly reacts to the tease (*oh shut up*) before returning to the initial topic of discussion, viz. his tutoring work. In other words, staged communicative acts typically involve the negotiation of a joint pretense, much in line with Clark's view on language as a joint activity.

When reviewing the literature on irony, sarcasm and teasing as staged communicative acts, it is apparent that there is an imbalance in the way in which the phenomena have been approached. A significant body of studies in (cognitive) linguistics has focused on the way in which speakers make use of contextually available information to construe a locally relevant ironic meaning (see e.g. Barnden, 2017 for an explicit positioning on this point). Next to the speaker's perspective, studies in psycholinguistics have zoomed in on the question of how addressees cognitively process irony (and more generally figurative language), including the matter of simultaneous vs. sequential processing of different accessible meanings (see e.g. Katz, 2017 for an overview). Significantly less studied in comparison to the speaker's and addressee's perspective, is the level of interaction involved. In other words, the question arises how language users jointly manage such staged communicative acts, which require forms of (multimodal) negotiation on the part of all participants involved.

To the best of our knowledge, only few studies zoom in on the way in which participants recruit various resources to participate in this process. Arguably the most

studied aspect relating to the marking of humorous intent (speaker perspective) or interpretation (addressee perspective), is laughter. Studies such as Glenn (2003), Bryant (2011), Holt (2016) and others have shown that laughter by all parties involved may play a constitutive role in negotiating the layered nature of the communication. Concerning facial expressions in the broad sense, in an exploratory study by Gironzetti et al (2016), it was observed that in humorous utterances in interaction, both speakers and addressees seem to pay increasing attention to the mouth and eyes of their co-participants (measured using eye-tracking systems), which may be indicative of a search for nonverbal markers of (ironic) intent and understanding. Tabacaru & Lemmens (2014) present a corpus-based study of sarcasm in interaction, in which they show that raised eyebrows on the part of the speaker may serve as gestural triggers, guiding the addressees towards the intended sarcastic interpretation of the utterance. Building on this study, Tabacaru (2019) argues that head tilts and head nods may serve a similar function as gestural triggers for sarcasm. Panzeri et al. (2019) provide experimental evidence that visual cues by a speaker, including facial expressions and bodily movements, enable interlocutors to recognize the ironic intent, even more so than the much-discussed ironic tone of voice as an acoustic cue. González-Fuente et al. (2015) show that speakers in two-party interactions change their gaze behavior more during ironic utterances than during non-ironic baseline utterances, which again may be indicative of speakers signaling a shift between viewpoints or layers of action. Brône & Oben (to appear) and Brône (2021) present similar evidence using eye-tracking data in three-party interactions. They show that both speakers and their addressees produce more gaze shifts in ironic utterances compared to non-ironic ones, and that addressees display more and longer moments of mutual gaze during ironic utterances, which may reflect a more complex grounding operation (i.e. addressees visually checking with their co-participants whether they have 'parsed' the utterance in a similar way). Needless to say, given the often exploratory nature of the studies mentioned here, there is a need for more systematic analyses that explore the (interplay between) various semiotic resources that interlocutors draw on to negotiate the staged communicative act.

1.1 The present study: focus and research questions

The present paper explicitly builds on some of the findings in the above-mentioned studies and aims to refine the existing evidence in a number of ways, including the granularity of analysis (zooming in on timing) and the interaction between resources. A temporal account with respect to timing is needed to gain a better insight into the interplay of multimodal marking of irony on the one hand, and the resources recruited for the micro-timing

of turn-taking in interaction on the other hand. Looking at the interaction between resources may be interesting with respect to the pragmatics of irony, with different resources potentially providing contradictory information or only very few elements 'hinting at' the intended ironic meaning. At the same time, this study may be viewed as a first step in uncovering 'multimodal Gestalts' (Mondada, 2014) for realizing the specific interactional project of irony.

Based on the literature overview sketched above, and on the gaps therein, we now describe our research questions. Existing research on which multimodal resources interlocutors recruit to participate in ironic interaction is scarce. Therefore our first aim is to replicate the findings in those studies, leading to the first research question:

Q1: Do speakers perform more gaze shifts (including gaze aversions), laughter and body movement (including gestures, head movements and shrugs) in ironic utterances, compared to non-ironic ones?

If employing multimodal resources has the potential to signal ironic intent, it might be fruitful for speakers to call in those resources at specific times during their conversational turn. At the level of eye gaze behavior, not only overall fixation counts or fixation durations might be relevant, but also more specific gaze patterns. In the present study we expect speakers to visually check whether or not their ironic intentions have been understood by their conversational partners. This visual grounding of ironic intention will mainly occur at the end of the utterances, i.e. after a speaker can expect a listener to have perceived that ironic intent. As such, we zoom in on the end of the utterance, resulting in Q2:

Q2: Do speakers perform more gaze shifts (towards, between and away from their conversational partners) at the end of their ironic utterances?

As opposed to verbal behavior, which is linear by nature, nonverbal behavior can be expressed in parallel at multiple levels. If the resources mentioned in Q1 appear to correlate with ironic utterances, a combination of those resources is also more likely to occur in ironic cases. This would be especially true for those cases where the negotiation of the ironic communicative act is more difficult or challenging. Since, to the best of our knowledge, this multimodal clustering has not received substantial scientific attention, we will explore the following research question:

Q3: Within ironic utterances, what clusters of gaze shifts, laughter and body movements arise?

Given the lack of studies zooming in on these multimodal packages of resources that could be linked to irony, we will not predict specific combinations of resources to be more frequent than others, but we will conduct an exploratory analysis of all possible combinations.

To make explicit how we tried to answer our research questions, we will first present the data collection and annotation procedure for each of the resources that we included in the study. The main findings of the study are presented in section 3, zooming in on the comparison of the ironic and non-ironic utterances in our dataset, as well as exploring the combinations of resources in ironic utterances and the temporal aspect of gaze behavior. In a final step, we discuss the results in light of recent findings in the literature.

2. Methods

2.1 Video corpus: participants and procedure

The video data used in this study were selected from the Insight Interaction Corpus (Brône & Oben, 2015), a corpus of spontaneous dyadic and triadic interactions in which participants wore head-mounted eye-trackers. All interactions are considered to be spontaneous, in the sense that they are not scripted. Participants were allowed maximal control over their contributions in the conversation: no roles, experimental tasks, time limits, etc. were imposed and participants were free to choose the length, the topic, the amount, the tone, etc. of their contributions to the conversation. The 'naturalness' of the interactions was aided by the fact that all members of a dyad or triad knew each other (well) prior to the recording. This allowed them to share a conversational and personal history, and to be more relaxed in the lab setting. All participants in the video corpus were students and native speakers of Dutch. The collection of this data set was in compliance with the local privacy and ethics committee at KU Leuven (under case number G-2015 02 173). Below we provide a basic description of the procedure and task design. For more detailed information, see Brône and Oben (2015) and Jehoul (2019).

After giving written consent to take part in the study, participants were invited to the recording room and seated in a triadic set up, enabling all participants to see each other with the same amount of effort. After participants were briefed on the recording session, and the eye-tracking glasses were calibrated, the experimenters left the recording room to allow for a conversational context as unobtrusive as possible. The recorded interactions consisted of two parts: the first part was a free conversation between all participants. For this interaction, participants were instructed to continue talking to each other as they did before the start of the recording (when they were waiting outside). In other words, they could talk freely about any topic they preferred. The second part of the interaction consisted of a brainstorm task. For this task, participants were asked to brainstorm about their ideal student bar and their ideal student house. Even though participants were instructed to talk about a given topic, this conversation too was highly interactive and not scripted. Participants were free to decide how rigorously they stuck to the brainstorm task. In practice, this often resulted in a brainstorm combined with free conversation

between participants (for instance about student bars that they went to).

For the purposes of the current study, we selected all brainstorms and spontaneous conversations from the triadic interactions. In total this amounted to 24 participants (3 male and 21 female) divided over 8 triads. Each triad took part in the free conversation and the brainstorm, which led to a sample of 16 videos (of 256 minutes of recordings in total). For the spontaneous interactions, a conversation lasted 21 minutes on average; for the brainstorms 18 minutes.

Recording equipment

All interactions were video-recorded using one external camera (Sony HDR-FX1000E, 25 frames per second, 720 × 576 pixels). All participants were also wearing Tobii Pro Glasses 2 (sample rate 50 Hz). These eye-tracking glasses are equipped with a scene camera and infra-red cameras. The infra-red cameras were used to capture the gaze focus of the participants, which was then mapped onto the scene cameras. All four camera perspectives were synchronized in one Quadvid to facilitate data analysis. A screenshot of the Quadvid can be found in Figure 1 below. In this figure, the top left camera shows the perspective of participant 1 in the picture in the bottom right. The top right picture shows the perspective from participant 2 and the bottom left camera shows the perspective from participant 3.

2.2 Data analysis

Annotation

As Gibbs (2000) notes, the diversity of language use in interaction, including various forms of non-literal language, makes it unappealing to compare ironic with general non-ironic utterances. To facilitate comparison of ironic with non-ironic utterances, we chose to zoom in on utterances that involve a positioning on an evaluative scale. In other words, both the ironic and non-ironic utterances in the dataset for this study can be categorized as *stance acts* (Du Bois, 2007). Many others have noted that irony can be seen as an evaluative phenomenon (e.g. Giora, 1995; Partington, 2007; Sperber & Wilson, 1986). See for example Burgers et al. (2011) for a more detailed account of such an approach to irony. Comparing ironic and non-ironic stance acts enables us to investigate which aspects of multimodal irony can be related to its highly intersubjective nature, beyond the evaluative function of irony.

The video data were annotated for *speech, irony, stance, gaze, laughter, head movement, shoulder shrugs, hand gestures and body repositionings*, using ELAN (Wittenburg et al., 2006) as an annotation tool. In our choice of multimodal resources, we took both evaluative and intersubjective aspects of irony into account.

The analysis of gaze behavior is inspired by previous studies (Brône, 2021; Brône & Oben, to appear; Giron-

Figure 1. Screenshot of Quadvid recording from the Insight Interaction Corpus.

zetti et al., 2016; González-Fuente et al., 2015). These studies demonstrated the salience of eye movements during humorous and ironic interactions, both in signaling the layered nature of the utterance, and in negotiating mutual understanding with other participants. The use of eye-tracking allows us to examine gaze behavior during ironic interaction in more detail.

Previous research (e.g. Brône et al., 2017; Holler & Kendrick, 2015) has demonstrated that gaze behavior is a temporally sensitive phenomenon. Therefore, we zoom in on gaze behavior at the end of ironic segments in addition to investigating gaze at the level of the whole utterance. The end of an utterance often marks a turn-transition relevant place (TRP), a hinge moment that is used by all interlocutors in a conversation to check mutual understanding (Goodwin & Goodwin, 1986; Sweetser & Stec, 2016). Therefore, we expect that gaze behavior will be especially salient during the last 1000 msec of the ironic utterance.

The use of head movements and shoulder shrugs in stance-taking is well documented (e.g. Debras & Cienki, 2012; Jehoul et al., 2017). The use of head tilts has been described as a "gestural trigger" for the use of irony (Tabacaru, 2019). Shoulder shrugs in ironic interaction have (to our knowledge) not been investigated systematically.

Hand gestures have not yet been systematically considered in the study of irony (with the exception of González-Fuente et al., 2015), because they are believed to convey more referential meaning, instead of pragmatic meaning (Tabacaru, 2019). However, pragmatic and interactive functions of gesture have been well established (Lopez-Ozieblo, 2020; Müller & Posner, 2004; Streeck,

1994), and therefore we explore the role of hand gestures in the expression of irony.

Finally, during the annotation of multimodal behavior, it appeared that participants displayed many body repositionings during ironic utterances, such as crossing or uncrossing the legs, folding the arms, etc. We therefore systematically annotated the presence of body repositionings during both ironic and non-ironic utterances.

As this study focuses on the use of nonlinguistic resources in irony, we consciously did not take into account classic paralinguistic features such as prosody.

For some of the variables described above (i.e. *laughter, gaze behavior, hand gestures* and *body repositionings*), the annotation process does not involve training or interpretation and is purely based on unambiguous, visual perception. These annotation processes are not subjected to a formal inter-coder agreement test. However, for those parameters that leave more room for interpretation (*irony, head movement type*), it is important to check the consistency and reliability of annotations. Therefore, we accounted for inter-rater reliability for these variables. In what follows, we describe the annotation process for all separate variables in more detail.

Irony

Previous research identified ironic utterances on many different levels, such as the sentence level (González-Fuente et al., 2015), clause level (Burgers et al., 2011) or utterance level (Bryant, 2011; Gibbs, 2000). In the current study, speech was transcribed using the GAT 2 transcription norms (Selting et al., 2009), at the level of the intonation unit. Segmenting spoken language into intonation

units, defined as segments with a coherent intonation contour corresponding with a new idea unit (Chafe, 1994), is a common practice in different types of interaction analysis. Irony was then annotated following the guidelines of Gibbs (2000). These guidelines were chosen because they have been used in earlier studies on the use of irony in interaction (e.g. Bryant, 2011; González-Fuente, 2015, Brône & Oben, to appear) and they provide a broad perspective on irony in interaction. In Gibbs (2000), irony is classified into five categories:

1. Jocularity: speakers tease one another in a humorous way;
2. Sarcasm: speakers speak positively to convey a more negative intent;
3. Rhetorical question: speakers literally ask a question that implies either a critical or a humorous assertion;
4. Hyperbole: speakers express nonliteral meaning by exaggerating the reality of the situation;
5. Understatement: speakers state far less than is obviously the case.

Each of these forms of irony "minimally reflects the idea of a speaker providing some contrast between expectation and reality" (Gibbs, 2000, p. 13).

In order to facilitate the comparison of ironic with non-ironic stance acts, we selected ironic segments that fall on an evaluative scale. For instance, one participant spoke about a time that she went to a camp, and slept in a house in the forest that had no lock. One of her co-participants reacts and utters (2):

(2) Ja da's nie eng (lacht)
    Yeah that's not scary (laughs)

    'k zou echt heel gerust slapen die nacht, amai
    I would really sleep very comfortably that night, oh boy

These two segments can be interpreted as ironic because the speaker provides a contrast between expectation and reality (i.e. the speaker pretends that the situation is not scary, whereas in reality it is). The segment is also a stance act, because the speaker indicates her affective position with respect to the missing lock in the house. Lastly, this affective stance can be put on a scale, on which the stance can be compared to other stances (such as that it is really okay to sleep in a house without a lock).

In another conversation, one participant (Amber) is telling two others about a fantasy she has in which she wants to push a cyclist off his bike, while he is cycling. One of the others (Lena) reacts and utters (3):

(3) (lacht) sadistje
    (laughs) little sadist

Although this example is a not a straightforward reversal of evaluation, it can still be considered ironic, as it reflects a contrast between expectation and reality. The inten-

tion of Lena was not to call her friend a person who takes pleasure in inflicting pain, punishment, or humiliation on others, but represents a teasing comment on the narration of the first participant. The segment is also a scalar evaluation of Amber, which can be compared to other evaluations (such as that this is a completely normal thought).

Non-ironic stance

The ironic segments in our dataset are compared to non-ironic evaluative segments. We placed evaluative segments in our dataset on a scale, ranging from a positive to negative evaluation. Specifically, we annotated explicit scalar evaluative segments such as (4), in which the participants are talking about the house in the forest without a lock, a bit further on in the conversation of utterance (2). The same co-participant reacts to this story and utters (4):

(4) Da's pas eng
    Now that is scary

This segment is an explicit evaluation (i.e. *this situation is scary*) and as such it can be placed on a scale and compared with other evaluations (like *this situation is not that scary*). Similarly, in (5) a participant reacts to a story of his friend, who explains that she usually discusses the rate of a taxi ride in advance with the driver.

(5) Ja da's wel slim
    Yeah that's smart

The speaker evaluates the segment of his friend, and his evaluation can be contrasted with other evaluations (like *that's stupid*).

Using the method described above, we annotated 123 scalar ironic cases in our data set. Three videos were excluded because they did not contain any case of scalar irony. For every ironic segment by a speaker in our corpus, we selected a non-ironic scalar evaluative segment by the same speaker, resulting in a total of 246 cases. An inter-rater-reliability test was conducted to ensure a proper selection of ironic and non-ironic cases. A second trained coder annotated roughly half (118 cases) of the selected data, and annotated whether a segment was "ironic" or "non-ironic". Agreement between the two raters was 91.5% (Cohen's Kappa = 0.831), providing us with confidence that the ironic cases were annotated correctly.

Gaze

Gaze of the interlocutors was annotated for the ironic and evaluative segments. To prevent influence from the speech in annotating gaze, this was done with the sound turned off. The areas of interest were defined as 1) the face of an interlocutor; 2) the body of an interlocutor; 3) background. Whenever a participant shifted their gaze to a new area of interest, a new gaze fixation was annotated. We set the minimum fixation duration at 120 msec,

in accordance with earlier research (Brône et al., 2017; Gullberg & Kita, 2009).

For the purposes of the current study, we measured gaze behavior of the speaker. Gaze aversions are defined as a gaze away from one of the listeners and toward the background. Speaker-to-listener gaze was defined as all fixations from the speaker on the face of either of the listeners. Mutual gaze was defined as the overlap in which a speaker looks at one of the listeners while that same listener looks at the speaker. Lastly, a gaze shift was defined as a shift to a new area of interest (either a participant or the background).

Gaze behavior during spontaneous social interaction is a temporally sensitive phenomenon. To explore the temporality of gaze behavior accompanying ironic versus non-ironic segments in more detail, we also annotated gaze behavior by the speaker in the last 1000 msec of a segment. During this timeframe, we measured to what extent there were gaze shifts from or towards the background (henceforth BG-shifts), between the two co-participants (henceforth CO-shifts) or no gaze shift at all (no-shifts). We also annotated the gaze state of the speaker at the end of the segment (i.e. gaze at a co-participant, or gaze at the background).

Laughter

The presence of laughter by the speaker during or immediately after the ironic and evaluative segments was annotated, at most 500 msec after the end of the segment.

Head movements

The presence of head movements during or immediately after the ironic and non-ironic evaluative segments was annotated. Head movements were divided into three categories: tilts (head movement where the ear moves toward the shoulder, left, right or both sequentially); nods (up- and downward movement of the nose) and shakes (left- and rightward movement of the nose). Head movements that did not belong to either of these categories (for instance when the head moves backward without any lateral movement), were categorized as "other". The presence of head movements was coded by one annotator. Head movement categories were then independently coded by two annotators, and cases of disagreement were discussed until agreement was reached for all items.

Shoulder shrugs

The presence of shoulder shrugs was annotated for all segments. A shoulder shrug could consist of either the left, the right, or both shoulders moving upwards.

Body repositioning

Body repositionings were defined as movements by a participant in which the major bodily articulators perform a body adjustment. Cases in which participants (un)crossed

their arms or legs or changed their seating position on the chair (e.g. shifting from 'hanging' to sitting upright, or swaying from a torso orientation leftwards to rightwards), were annotated as a binary variable (that is, there either is or is no body repositioning during a segment).

Hand gestures

Hand gestures were identified as communicative movements by the hand(s). Only hand and arm movements that were not self-adaptors were considered. In the current study, gestures were not further categorized into formal characteristics or functional gesture types.

3. Results

In this paper we investigated which semiotic resources are recruited by speakers in the expression of irony in spontaneous interaction, as well as how those resources interact. In the first part of this section, we describe the distribution of these semiotic resources in ironic and non-ironic segments. For this analysis, the statistical analysis software R (R Core Team, 2021) was used. In the second part of this section, we take a more qualitative approach and explore how these resources interact, and how their variation in use can be explained.

3.1 Recruitment of individual resources in expressing irony

In the current corpus of 16 conversations, we found 123 scalar ironic segments, enough for a small quantitative study. Three of the interactions did not contain any ironic segment, and were thus excluded from this analysis. We then annotated laughter, gaze behavior, head movements, shoulder movements, hand gestures and body repositionings for both ironic and non-ironic segments. To provide a first insight in the distribution of bodily resources, we added up those resources that can be reduced to a binary score (*laughter, CO-shift at end of segment, head movement, shoulder shrug, hand gesture, body repositioning*). In Table 1 below, the number of resources involved in the expression of ironic and non-ironic stance acts is displayed. A Wilcoxon test for independent samples was performed to explore differences between the two conditions in the recruitment of multimodal resources. The results of this

Table 1. Number of resources involved in ironic and non-ironic segments.

Nr of resources involved	Frequency	
Condition	Ironic	Non-ironic
0	13	51
1	33	45
2	52	20
3+	26	6

test showed that there is a larger number of resources accompanying ironic ($Mdn = 2$) compared to non-ironic ones ($Mdn = 1$), a significant difference (Ws (241) = 11209.5, p < 0.0001)[1], with a moderate effect size (Hedges $g_s = 0.566$)[2].

In what follows, we will describe the contributions of all separate resources to the expression of irony in more detail.

Laughter

An obvious marker of irony can be found in laughter. We investigated laughter by the speaker accompanying ironic and non-ironic segments. There was more laughter in ironic compared to non-ironic segments. Out of all 123 ironic segments, 62 were accompanied by laughter, compared to 8 out of 123 for non-ironic segments. A chi-square test showed that this difference was significant (χ^2 (1) = 56.089, p < 0.0001), and the effect was moderate (Cramer's $V = 0.477$)[3].

Gaze

With respect to gaze behavior across the segment, we were interested in four variables: gaze aversions, speaker-to-listener gaze, mutual gaze, and gaze shifts. There does not seem to be a difference in the counts of mutual gaze between ironic and non-ironic segments, which was shown by a Wilcoxon test for independent samples ($Ws = 4297$, p = 0.330, Hedges' $g = 0.006$). Similarly, we did not find a significant difference in the amount of speaker-to-listener gazes ($Ws = 7729.5$, p = 0.288, $g_s = 0.146$). We also did not observe a significant difference in the amount of gaze aversions by the speaker ($Ws = 7787$, p = 0.122, $g_s = 0.224$) or in the amount of gaze shifts ($Ws = 8029.5$, p = 0.101, $g_s = 0.245$).

Regarding gaze duration, again no differences arise between the two conditions with respect the duration of mutual gaze ($Ws = 4111.5$, p = 0.175, $g_s = -0.095$), speaker-to-listener gaze duration ($Ws = 7616$, p = 0.436, $g_s = .034$) or the duration of gaze aversions ($Ws = 7669$, p = 0.095, $g_s = 0.095$).[4] Descriptive statistics of all gaze variables are listed in Table 4.

As mentioned above, we also investigated gaze behavior during the last 1000 msec of the segment in more detail. Within this timeframe, we analyzed the presence of gaze shifts and the end state of the gaze by the speaker. We found that a different gaze pattern emerges at the end of ironic segments compared to non-ironic segments. At the end of ironic segments, participants more often show gaze shifts from and to the background (30 out of 123), as well as more gaze shifts between listeners (38 out of 123), and less continuous gaze towards one AOI (or "no-shift") (46 out of 123), compared to non-ironic segments (22, 27 and 66 out of 123 respectively). This difference was significant in a Chi-square test (χ^2 (2)= 10.893, p = 0.004, Cramer's $V = 0.171$). As for the gaze state at the end of the segment, no different pattern emerges for ironic compared to non-ironic segments (χ^2 (1) = 0.304, p = 0.582, Cramer's $V = 0.046$). In both conditions, participants look at their interlocutors more often than they look away.

Head movements

The presence of head movements was also annotated for ironic and non-ironic segments. Here we found that ironic segments are more often accompanied by head movements (63 out of 123) compared to non-ironic segments (40 out of 123) (χ^2 (1) = 8.812, p = 0.003), which

Table 2. Frequencies of head movement types during ironic and non-ironic segments.

Head movement type	Ironic	Non-ironic
Nod	24	23
Shake	17	14
Tilt	24	10
Other	8	2

is a moderate effect ($V = 0.200$). We then explored head movements in more detail, by looking at differences in head movement types. The frequencies of head movement types per condition are listed in Table 2.[5]

From this more granular analysis, it appears that a difference in the amount of head movements can mostly be attributed to the increase of head tilts in ironic segments compared to non-ironic segments.

---

1   As the data in our dataset are not normally distributed, we reported Medians and Range instead of Means and Standard Deviations, and conducted non-parametric tests that are equivalent to parametric t-tests, such as the Wilcoxon Rank Sum test.

2   Hedges G is a standardized effect size that can be used to interpret findings from non-parametric statistical tests such as Wilcoxon Independent Samples test. For more detailed information about the use and interpretation of Hedges G, see Delacre et al. (2021).

3   Cramer's V is a standardized measure of effect size for Chi-square tests, and can be interpreted following Cohen's (1988) guidelines.

4   There was no significant difference in the total duration of ironic versus non-ironic utterances. (W = 7526.5, p = 0.9464).

5   Some utterances were accompanied by multiple head movements (for instance both a tilt and a nod). For these utterances we counted each head movement as a separate type, leading to a small discrepancy between the total number of utterances that were accompanied by head movement, and the total frequencies of different head movement types.

Shoulder shrugs

The annotation of shoulder shrugs showed that there were very few cases in this dataset. There were no differences between shoulder shrugs in ironic (10 out of 123) compared to non-ironic segments (9 out of 123), (χ^2 (1) = 0, p = 1, V = 0.014).

Body repositioning

Sometimes participants re-adjusted themselves during a segment. When counting this systematically, it appeared that there were more body repositionings in ironic segments (21 out of 123) compared to non-ironic segments (6 out of 123) (χ^2 (1) = 8.031, p = 0.005), a small effect (V = .194).

Hand gestures

We investigated the overall presence of gesture in ironic and non-ironic segments, disregarding different gesture types or form-based characteristics. The results show that 34 out of 123 of ironic segments are accompanied by gesture, compared to 23 out of 123 for non-ironic segments. This difference was not significant in a Chi-square test (χ^2 (1) = 2.284, p = 0.131, V = 0.106).

All results are summarized in Tables 3 and 4. These Tables show that ironic segments are accompanied by relatively more laughter, head movements, body repositionings, and gaze shifts, as well as gazes from and to the co-participants at the end of the segment. So far, we have focused on the occurrence (or lack thereof) of a variety of individual resources in ironic and non-ironic cases. Below, we move to a more encompassing take on the phenomenon by analyzing the co-occurrence of those resources in ironic interaction. Those potential multimodal packages

Table 3. Total counts of resources involved in the expression of ironic and non-ironic segments, and proportion of occurrence within the condition.

Variable	Condition	Raw frequency	Proportion
Laughter *	Ironic	62	0.504
	Non-ironic	8	0.065
Hand gesture	Ironic	34	0.276
	Non-ironic	23	0.187
Head movements *	Ironic	64	0.520
	Non-ironic	40	0.325
Body repositioning *	Ironic	21	0.171
	Non-ironic	6	0.049
Shoulder shrug	Ironic	9	0.073
	Non-ironic	8	0.065
Gaze shift end of segment *	Ironic	68	0.553
	Non-ironic	49	0.398
Gaze at co-participant end of segment	Ironic	85	0.691
	Non-ironic	81	0.659

NB. Resources of which the frequencies of occurrence differ significantly between conditions, are marked with an asterisk.

Table 4. Descriptive statistics for gaze behavior during ironic and non-ironic segments.

Variable	Condition	Median (IQR) duration		Median (IQR) count	
Mutual gaze	Ironic	619.0	[50–1002]	1	[1–1
	Non-ironic	609.5	[125.2–1116]	1	[1–1.25]
Speaker-to-listener gaze	Ironic	853.5	[372.5–1344.2]	1	[1–2]
	Non-ironic	767.0	[141.2–1425.5]	1	[1–2]
Gaze aversion	Ironic	0	[0–274.2]	0	[0–1]
	Non-ironic	0	[0–0]	0	[0–0]
Gaze shift	Ironic	–		1	[0–6]
	Non-ironic	–		1	[0–6]

will first be tackled from a quantitative perspective, and subsequently discussed further in a qualitative analysis.

3.2 Multimodal packages

As mentioned in the introduction, the current paper aims to refine the existing evidence for embodied communication of irony by investigating how multiple resources are employed in spontaneous interaction. In this section, we answer two questions related to this overarching goal. Firstly, we explore which resources co-occur systematically in the expression of ironic stance, in the form of "multimodal packages". Secondly, we explore what factors might drive variation in the markedness (i.e. the amount of resources involved in the expression) of an ironic segment.

Co-occurrence of resources

To explore the co-occurrence of different markers of irony quantitatively, we calculated Kendall's Tau correlations between all of the resources under scrutiny in the current paper: *gaze behavior* (*mutual gaze, speaker-to-listener gaze, gaze aversions, gaze shifts, CO-shifts at the end of a segment*), *laughter, head movements, shoulder shrugs, body repositionings,* and *hand gestures*. We did this only for the group of ironic segments. Below we discuss those combinations of resources for which a significant correlation was found.

As a first result, this analysis shows a correlation between the presence of laughter, and different measures of gaze behavior. That is, in ironic segments where there is laughter, there are also more gaze shifts (T = 0.218, p = 0.009), more instances of mutual gaze (T = 0.177, p = 0.040) and speaker-to-listener gaze (T = 0.219, p = 0.010), as well as gaze aversions (T = 0.243, p = 0.007). Note that these different measures of gaze behavior are also correlated. Interestingly, there was no correlation between the presence of CO-shifts at the end of a segment, and the presence of laughter.

Let us now take a look at some examples from the dataset to examine the relation between laughter and gaze in more detail. In the fragment below, three participants (Anna, Ella and Seb) are talking about the father of Anna,

one of the participants. Anna explains that her father is an autodidact and invests much time in self-study. He speaks better French than Anna, and he can translate Latin texts even though he never studied Latin in school. In lines 1 and 2 of the fragment below, Ella comments on this, showing her non-ironic stance. In line 3, Anna adds to this by saying *makes you cry, right*. As this does not really make Anna cry, this segment can be interpreted as ironic. Anna takes on a pretense role in which she is so upset about her father's knowledge that it makes her cry. Her laughter directly following the segment marks the non-seriousness of the segment.

In order to represent the distribution of visual attention of all participants, we use a score-like representation for their gaze behavior in co-occurrence with the transcript lines. The symbols in the score represent the gaze target at each point in time (e.g. in line 1, the current speaker Ella shows sustained gaze to the background (bg), whereas Anna shifts her gaze from the background to Ella towards the end of the short segment. The third participant, Seb, gazes at Anna for the entire duration of the segment). By representing the gaze direction of all three participants in this way, we get a detailed picture of the distribution of visual attention at each point in time, related to the segments being produced. When looking at the gaze score for the ironic segment in fragment 6, it is clear that Anna shifts her gaze between the background and her co-participants. In this example, both the use of gaze shifts and laughter can be argued to form interactive devices, inviting her audience to align with her and join her in this pretense mode.

```
(6)        (zucht) maf
   1 Ella   (sighs) nuts
     Anna   bg------Ella
     Ella   bg-----------
     Seb    Anna-------
            Echt [maf;]
   2 Ella   really [nuts;]
            [is we]nen he (lacht)
→  3 Anna   [makes you cr]y right, (laughs)
     Anna   Ella------bg-----Seb---Ella-------------------
     Ella   bg----------------------------------Anna----
     Seb    Anna-------------------------------------------
```

In another example, shown in fragment 7 below, three participants are brainstorming about their ideal student bar. The participants (Leah, Karen and Mia) suggest that it would be nice to serve some food first, like nachos. Leah jokingly says "for free", and laughs. This marks the beginning of the transcript, in line 1. Line 1 can be interpreted as ironic since, in reality, student bars would most likely not serve food for free. By laughing, Leah marks this segment as playful and elicits a response from her addressees. In line 2, Mia returns to a serious statement, and starts with another suggestion. She uses a Palm Up Open Hand (PUOH) gesture. While she gestures, Leah shifts her gaze towards Mia and takes the turn. In line 3, Leah

expands on her initial statement, also using gesture (a Palm Down Open Hand gesture, PDOH) and still looking at Mia. Mia then shifts her gaze toward Leah. Again, Leah laughs and this time Mia responds by laughing (in line 5), aligning her gesture with Mia's PDOH gestures and repeating Mia's line "everything for free". In this example, both the use of laughter and mutual gaze contribute to the joint act of irony, by eliciting a reaction from the addressees and involving them in the staged communicative act.

```
(7)          gratis;
   1 Leah   for free ;
            ((Leah laughs))
     Leah   Karen----------
     Karen  Leah-----------
     Mia    bg--------------
            +en wat mis[schien ook wel fijn is ;+
→  2 Mia    and what might also be nice ;
            +Mia gestures PUOH-----------------+
     Leah   Karen----Mia-----------------------------
     Karen  Leah----ges Mia-------Leah-----------
     Mia    bg------------------Leah----------------
            *[[alles] gratis ; *
   3 Leah   [[everything] for free ;
            *Leah gestures PDOH-----*
   4 Karen  [a:h -]
            ((Leah laughs))
            <<lachend> ja (.) alles *$gra*$tis ja .>
   5 Mia    <<laughing> yeah (.) everything for free
            yes.>
            *Mia gestures PDOH*
            $Mia shakes head$
     Leah   Mia----------------------bg------Karen-
     Karen  Leah------------------Mia----Leah------
     Mia    Leah-----------------bg---------------
```

Turning to other systematic co-occurrences of resources in the expression of irony, we found that the frequency of mutual gaze was moderately correlated with the presence of hand gestures during ironic segments (T = 0.201, p = 0.020). An example of this can be found in the fragment above. In this example both Leah and Mia use gesture accompanying their ironic segments. Specifically in line 3 where Leah repeats her initial ironic segment, Mia and her display mutual gaze. After this, Mia repeats Leah's gesture as well as her verbal expression, joining her in the pretense.

Another small but significant negative correlation, was found between the presence of gesture and shoulder shrugs (T = -0.184, p = 0.042). Please note that this result should be interpreted with caution as there are only few attestations of shoulder shrugs in this dataset. Again a negative correlation is found between head movements and body repositionings, demonstrating fewer head movements in segments with body repositionings (T = -0.249, p = 0.006).

Lastly, as mentioned above, several measures of gaze behavior were correlated. The frequency of gaze shifts

correlated with the presence of all other gaze variables, including CO-shifts at the end of the segments. In the same line of reasoning, the number of mutual gazes was strongly correlated with the number of speaker-to-listener gazes. Finally, the number of gaze aversions was also correlated with the number of speaker-to-listener gazes.

Summarizing these findings, we see a co-occurrence of laughter and various measures of gaze behavior throughout the segment. The presence of gesture is positively correlated with the presence of mutual gaze, and negatively correlated with the presence of shoulder shrugs. Head movements and body repositionings are also found to be correlated. Finally, various measures of gaze behavior are correlated.

Variation in the multimodal expression of irony

Although some resources, such as the use of gaze shifts and laughter, co-occur more systematically in the expression of irony, the current dataset also shows quite some variation in the multimodal "marking" of irony. In this section, we explore the question of why some ironic expressions are marked more than others.

A first obvious factor that might play a role here, is the duration of the segment: the longer a segment, the more opportunities a speaker has to use multiple resources. And indeed, the amount of resources involved in the expression of irony is correlated with the duration of ironic segments ($r(121) = 0.279$, 95% CI [0.107;0.434], $p = 0.002$). Longer segments are accompanied by more resources. Interestingly, this correlation between duration of a segment and number of resources does not arise for non-ironic segments ($r(120) = -0.028$, 95% CI [-0.204;0.151], $p = 0.763$). Furthermore, there is no difference in duration of ironic versus non-ironic segments. As a consequence, duration alone cannot explain the amount of resources used. In other words, the amount of resources we observe cannot be solely attributed to the chance-level expectation that arises from the fact that longer segments are by default more likely to contain more of the resources under scrutiny in the present paper.

Another factor that might be at play, is the difficulty for conversational partners to understand an ironic segment as non-serious. One possibility is that ironic segments that are difficult to understand, for instance due to a lack of context, are marked using multiple bodily resources in order to facilitate comprehension for the listeners. However, a qualitative analysis shows that this does not seem to be the case. On the contrary, ironic segments that are part of an ironic sequence of some sort (i.e. that are immediately preceded or followed by another ironic segment by the same or another speaker), frequently occur with three or more bodily resources. This conflicts

with the expectation that segments that are more straightforwardly recognizable as being ironic (i.e. because they are part of a sequence), require less multimodal marking. Of all ironic segments marked by three or more resources, 19 out of 26 were preceded or followed by another ironic remark within two turns, compared to 6 out of 13 segments that were not marked by any resources.

The following example in fragment 8 below provides an illustration of this embeddedness. Three participants (Paul, Gabriëlle, and Mara) are speaking about Mara's sister, who is a very good singer. Mara explains that her sister sings at home, but she never performs at other places. Paul then comments that she has a hidden talent. This marks the beginning of the fragment. Gabriëlle mumbles that she has no hidden talent, the first ironic segment. Mara then responds to her in 3, saying that "it all radiates from her". This is an obvious non-serious teasing. Mara and Gabriëlle are friends, and this sarcastic remark doesn't appear to be meant as harmful. Instead, both Gabriëlle and Mara seem to adopt a pretense reading about the state of Gabriëlle's talents. All of the participants mark their mutual understanding of this pretense by responding to the segment by laughing. Segment 3 is marked in a number of ways. Mara uses a metaphoric iconic gesture that represents the radiation of Mara's talent (see also Figure 2). Secondly, she shifts her gaze from Gabriëlle towards Paul at the end of the segment. Third, the speaker laughs both before and immediately after the segment.

```
(8)              <<mompelt>> Ik heb geen verborgen talent. >
→  1 G           <<mumbles> I have no hidden talent . >
     P           M-------------------G----------------
     M           bg---------------G-------------------
     G           M------------------------------------
                 ((G laughs))
                     ((P laughs))
                         ((M laughs))
     2 P         (inau[dible)
                 [bij u straalt het er gewoon *allemaal [af]*.⁶
→  3 M           In your case it just all radiates from you.
                                     *gestures------*
                                         [G laughs]
     P           G------------------------G----------------
     M           G--------------------------------P---------
     G           bg---------------------M-------------------
                 ((M laughs))
                 ((P laughs))
```

Another example of a highly marked ironic segment is found in the transcript concerning fragment 7, the brainstorm about a student bar. When Mia says "yeah everything for free", she uses hand gestures, head shake and laughter accompanying this segment. However, these

---

6 In the Dutch original utterance the word order is different, so that the iconic gesture is produced *after* the word "radiates", during "all off". For a complete understanding, please see the glossed translation: *Bij* – With, *U* –you, *straalt* – radiates, *het* – it, *er* – there, *Gewoon* – Just, *allemaal* – all, *af* – off.

"In your case it just all radiates from you."

Figure 2. Iconic gesture accompanying "all from you", both hands moving outwards to represent radiation.

markers are not necessary for the interpretation of the ironic segment. From the lexical content only, it becomes clear that this segment is not meant as a serious statement. A student bar cannot serve free food. Similar to the "hidden talent" example, this segment is also embedded in a short ironic sequence in which all participants contribute to the joint pretense that the food (or even everything) is for free, adding to the obviousness of the irony in this segment.

Summarizing this exploration of the use of multiple resources in the expression of irony, we find that the use of laughter systematically co-occurs with the use of marked gaze behavior and that the use of head movements is negatively associated with the presence of body repositionings. Furthermore, variation in the amount of resources involved in the expression of irony can be partially explained by the duration of the segment. We also hypothesize that the embeddedness in an ironic sequence plays a role. We now move on to the discussion of our findings in light of the existing literature, and propose some ways to move forward.

4. Discussion

Research in a variety of disciplines, including cognitive science, linguistics, literary analysis and philosophy, has investigated how the notion of viewpoint permeates all products of human cognition, from mundane talk-in-interaction over intricate narratives to physical works of art (Dancygier et al., 2016; Dancygier & Sweetser, 2012; McIntyre, 2006; Tobin, 2018; Vandelanotte, 2017, and many others). This is also the case for verbal irony, where a large body of research in cognitive linguistics has focused on the way in which speakers make use of contextually available information to construe a locally relevant ironic meaning. However, the ways in which language users construe and negotiate such staged communicative acts in an interactional context, remain largely underexplored. In this study, we investigated how speakers in spontaneous three-party interaction recruit bodily resources in the communication of irony. This paper addresses three research questions: Q1) Do speakers perform more gaze shifts (including gaze aversions),

laughter and body movement (including gestures, head movements and shrugs) in ironic utterances, compared to non-ironic ones; Q2) Do speakers perform more gaze shifts (towards, between and away from their conversational partners) at the end of their ironic utterances; Q3) Within ironic utterances, what clusters of gaze shifts, laughter and body movements arise?

We took a broad approach and investigated the role of laughter, gaze behavior, head movements, shoulder shrugs, body repositionings and hand gestures. These variables have all been found to play a relevant role in the expression of irony, and stance more broadly (Attardo et al., 2013; Brône & Oben, to appear; Debras & Cienki, 2012; Gironzetti et al., 2016; González-Fuente et al., 2015). We found that, on average, participants used more of these bodily resources in the expression of ironic stance, compared to non-ironic stance. These results underline the importance of multimodality in the communication of irony in face-to-face interaction. The involvement of bodily resources was mainly manifested in the use of laughter, gaze behavior (i.e. CO-shifts in the last 1000 msec of the segment), head movements and body repositionings.

The finding that speakers use more laughter accompanying ironic compared to non-ironic segments, is in line with previous research on this topic, which demonstrates that laughter can signal ironic intent of a speaker, or mark segments as playful (Bryant, 2011; Holt, 2016). Laughter, and especially antiphonal laughter, also has an interactive function as a marker of stance alignment (Attardo et al., 2013; Bryant, 2011).

Zooming in on the end of the segment, speakers display more gaze shifts both between their addressees and to and from the background. These findings are in line with a view of irony as a highly intersubjective process, during which it is necessary for the speakers to check whether their interlocutors have parsed their segments as intended. We did not, however, observe an increase in mutual gaze between speakers and their addressees at the level of the entire ironic segment, in contrast to what others have observed (Brône & Oben, 2021 observed a higher amount of gaze between addressees in ironic v. non-ironic segments). Neither did we, as yet others have observed (González-Fuente et al., 2015; Williams et al., 2009), find an increase in gaze aversion of speakers throughout the segment. To further investigate the precise nature of gaze behavior during ironic interaction, it is necessary to take into account the turn structure, as there is a large body of research showing robust patterns of gaze in social interaction (e.g. Brône et al., 2017).

Head tilts, too, have been noted to function as irony markers (Tabacaru, 2019), but have been more prominent in relation to stance (Debras & Cienki, 2012; Jehoul et al., 2017). The presence of more head tilts in ironic compared to non-ironic stance-taking, is in line with a function of head tilts as displaying the layered viewpoints in interaction, much like has been described for enactments of third parties (Debras & Cienki, 2012).

Lastly, more body repositionings were found in ironic compared to non-ironic segments. Follow-up research will have to point out whether this effect will hold with a larger sample size, and investigate the precise nature of this effect. Body repositioning in our annotation should in no case be confused with body partitioning that e.g. Debras (2015), Janzen (2019) or Sweetser & Stec (2016) link to shifts in viewpoint. For these authors body partitioning occurs when one part of a body (e.g. head and gaze direction) is associated with one viewpoint, whereas another part of the body (e.g. torso orientation) is simultaneously associated with another viewpoint. This allows speakers to simultaneously display two entities/viewpoints or perform a shift in entity/viewpoint while maintaining the current one. In our data, body repositioning bears more resemblance to fidgeting or self-adaptors (see e.g. Bressem & Ladewig, 2011, or Freedman, 1977). The question that still remains is how this seemingly non-intentional phenomenon, that appears to be low in carrying communicative information, can be relevant in producing or observing ironic segments.

Relating to the interaction between these resources, we found that the presence of laughter was associated with more marked gaze behavior: more gaze shifts, moments of mutual gaze, speaker-to-listener gaze, and gaze aversions. In a small qualitative analysis, we showed that the joint use of gaze and laughter can be used to invite addressees to join an ironic pretense in interaction. Interestingly, there was no correlation between the presence of laughter and gaze shifts at the end of the segment, speaking against the idea that their co-occurrence is based only on alignment-seeking with co-participants. Future research on the precise timing of both of these resources could shed more light on the nature of their interaction.

We also found that the presence of gesture was associated with more mutual gaze in ironic segments. An interesting follow-up would be to investigate what kind of gestures are associated with mutual gaze (for instance the use of pragmatic and interactive gestures, Bavelas et al., 1992; Lopez-Ozieblo, 2020), and how they relate to one another temporally.

There was a negative association between the presence of body shifts and head movements. One possible explanation is that body repositionings require the use of the whole body including the head, which is then not available to be used as a communicative marker. However, as discussed, more research is needed to determine the nature of body repositionings in ironic interaction in general, before we are able to draw any conclusions regarding its interaction with other communicative devices. The same holds for the negative association that was found between the presence of shoulder shrugs and gestures.

Finally, several measures of gaze behavior were correlated. The frequency of gaze shifts correlated with the presence of all other gaze variables. This is a straightforward finding, as the presence of gaze shifts automatically assumes the presence of more gaze fixations, be it gaze aversions, mutual gaze, or speaker-to-listener gaze. In the same line of reasoning, the number of mutual gazes was strongly correlated with the number of speaker-to-listener gazes. Finally, the number of gaze aversions was also correlated with the number of speaker-to-listener gazes. This finding can be attributed to the fact that for a fixation to be counted as a gaze aversion, there must always be a speaker-to-listener gaze prior to this fixation. Lastly, the number of gaze shifts was correlated with CO-shifts at the end of the segment. In ironic segments that display more gaze shifts throughout the segment, it seems logical that there will also be more gaze shifts at the end of the segment.

In investigating the variation between the amount of resources involved, we explored whether this variation could be due to potential processing difficulties. Previous research on comprehension of irony showed that visual cues by a speaker, including facial expressions and bodily movements, enable interlocutors to recognize the ironic intent, even more so than the much-discussed ironic tone of voice as an acoustic cue (Panzeri et al., 2019). Following this line of reasoning, we explored whether ironic segments that are marked more extensively are also more difficult to understand. However, in a qualitative observation of the segments that were marked with 3+ resources and with 0 resources, we found that this was not the case. Segments that were marked more extensively often occurred within a context of other ironic segments, presumably facilitating the interpretation of an ironic segment as such. Segments that were not marked at all, however, often occurred without such a context. This raises interesting questions concerning the function of embodied resources in communicating irony. In general, the distribution of ironic sequences in interaction remains understudied, and studies investigating this topic showed conflicting results. Gibbs (2000) found that in conversations among friends, around thirty percent of ironic segments was followed by another ironic segment. However, in their analysis of a large number of fieldnotes of interactions, Eisterhold et al. (2006) found that only 6% of ironic segments were followed by an ironic segment.

Another factor that impacted the degree of resources involved was the duration of the segment: longer ironic segments were accompanied by more resources than shorter ironic segments. However, it is unlikely that the duration of the segment was the only factor at play, as the duration was not correlated with the amount of resources involved in the expression of non-ironic stance. Therefore, follow-up studies investigating the embeddedness of ironic segments in so-called "ironic sequences", in which all participants in a conversation join in this staged communicative act, are necessary to disentangle the role of the body in communicating and negotiating irony.

One factor that was left unexplored with respect to variation in the resources involved, is individual differences between participants. Although the use of quantitative models that can take into account such individual differences, like linear mixed effects models (cf. Baayen et al., 2008) are now standard in linguistics, the small

sample size and explorative nature of the current study did not allow us to consider this.

Another limitation with regard to variation in bodily marking of irony, is that segments that were categorized as "non-marked" in our study might have had some form of bodily involvement that was not part of our analysis. The current study did not take into account the role of prosody, and the role of facial expressions. Within the limited scope of this study, we decided to zoom in on visual bodily behavior rather than paralinguistic features, which have already received a significant amount of scholarly attention. The reason for excluding facial expressions as a feature of visible bodily behavior is of a more practical nature: the participants in the recorded interactions are all wearing eye-tracking glasses, which cover essential features for the analysis of facial expressions (such as eyebrows) .

Finally, the current study focused on the behavior of the speaker within single segments. A qualitative exploration in the current study did bring forward the hypothesis that the embeddedness of ironic segments in an interactional context may play a crucial role in explaining the involvement of various bodily resources such as laughter and gaze behavior. This study can thus be interpreted as an invitation for a more systematic study of the role of interactional context in the multimodal expression of irony, taking into account behavior by all participants over longer stretches of ironic interaction.

5. Conclusion

In conclusion, this study investigated the multimodal expression of irony in interaction. Ironic utterances can be viewed as a form of staged communicative acts (Clark, 1996) which require forms of (multimodal) negotiation on the part of all participants involved. We showed that, during ironic utterances, speakers employ a range of resources, most notably head tilts, laughter, gaze shifts at the end of segments and body repositionings. Furthermore, the use of laughter and marked gaze behavior, as well as the use of gesture and gaze behavior, systematically co-occurs during ironic utterances. Together, the speaker recruits these bodily resources to bring the attention of the audience to the multiple layers at play in irony, and to invite the co-participants to join in this pretense. The results of the current study speak to a notion of irony as a jointly construed and negotiated form of pretense (cf. Brône, 2021; Brône & Oben, 2021), and put forward the question of how the interactional structure influences the use of multimodal resources in the expression of irony.

References

Attardo, S., Pickering, L., Lomotey, F., & Menjo, S. (2013). Multimodality in conversational humor. *Review of Cognitive Linguistics*, *11*(2), 402–416. https://doi.org/10.1075/rcl.11.2.12att

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barnden, J. (2017). Irony, pretense and fictively-elaborating hyperbole. In A. Athanasiadou & H. L. Colston (Eds.), *Irony in language use and communication* (pp. 145–178). John Benjamins Publishing Company. https://doi.org/10.1075/ftl.1.08bar

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, *15*(4), 469–489. https://doi.org/10.1080/01638539209544823

Bressem, J., & Ladewig, S. H. (2011). *Rethinking gesture phases: Articulatory features of gestural movement? 2011*(184), 53–91. https://doi.org/10.1515/semi.2011.022

Brône, G. (2008). Hyper- and misunderstanding in interactional humor. *Journal of Pragmatics*, *40*(12), 2027–2061. https://doi.org/10.1016/j.pragma.2008.04.011

Brône, G. (2021). The multimodal negotiation of irony and humor in interaction: On the role of eye gaze in joint pretense. In A. Soares da Silva (Ed.), *Figurative Thought and Language* (Vol. 11). John Benjamins Publishing Company, 109–136. https://doi.org/10.1075/ftl.11.04bro

Brône, G., & Oben, B. (2015). InSight Interaction: A multimodal and multifocal dialogue corpus. *Language Resources and Evaluation*, *49*(1), 195–214. https://doi.org/10.1007/s10579-014-9283-2

Brône, G., & Oben, B. (to appear). *Monitoring the pretence. Intersubjective grounding, gaze and irony*.

Brône, G., Oben, B., Jehoul, A., Vranjes, J., & Feyaerts, K. (2017). Eye gaze and viewpoint in multimodal interaction management. *Cognitive Linguistics*, *28*(3), 449–483. https://doi.org/10.1515/cog-2016-0119

Bryant, G. A. (2011). Verbal irony in the wild. *Pragmatics & Cognition*, *19*(2), 291–309. https://doi.org/10.1075/pc.19.2.06bry

Burgers, C., van Mulken, M., & Schellens, P. J. (2011). Finding Irony: An Introduction of the Verbal Irony Procedure (VIP). *Metaphor and Symbol*, *26*(3), 186–205. https://doi.org/10.1080/10926488.2011.583194

Chafe, W. (1994). *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press.

Clark, H. H. (1996). *Using Language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620539

Clark, H. H., & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of Experimental Psychology: General*, *113*(1), 121–126. https://doi.org/10.1037/0096-3445.113.1.121

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Coulson, S. (2005). Sarcasm and the Space Structuring Model. In S. Coulson & B. Lewandowska-Tomaszczyk (Eds.), *The Literal and Nonliteral in Language and Thought* (pp. 129–144). Peter Lang.

Dancygier, B. (Ed.). (2017). *The Cambridge Handbook of Cognitive Linguistics*. Cambridge University Press. https://doi.org/10.1017/9781316339732

Dancygier, B., Lu, W., & Verhagen, A. (Eds.). (2016). *Viewpoint and the fabric of meaning: Form and use of viewpoint tools across languages and modalities*. De Gruyter Mouton. https://doi.org/10.1515/9783110365467

Dancygier, B., & Sweetser, E. (2012). *Viewpoint in Language: A Multimodal Perspective*. Cambridge University Press. https://doi.org/10.1017/CBO9781139084727

Dancygier, B., & Vandelanotte, L. (2016). Discourse viewpoint as network. In B. Dancygier, W. Lu, & A. Verhagen (Eds.), *Viewpoint and*

the fabric of meaning: Form and use of viewpoint tools across languages and modalities (pp. 13–40). De Gruyter Mouton. https://doi.org/10.1515/9783110365467-003

Debras, C. (2015, September). Stance-Taking Functions of Multimodal Constructed Dialogue during Spoken Interaction. *Proceedings of Gesture and Speech in Interaction 4*. https://hal.parisnanterre.fr//hal-01640486

Debras, C., & Cienki, A. (2012). Some Uses of Head Tilts and Shoulder Shrugs during Human Interaction, and Their Relation to Stancetaking. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, 932–937. https://doi.org/10.1109/SocialCom-PASSAT.2012.136

Delacre, M., Lakens, D., Ley, C., Liu, L., & Leys, C. (2021). *Why Hedges' g*s based on the non-pooled standard deviation should be reported with Welch's t-test*. PsyArXiv. https://doi.org/10.31234/osf.io/tu6mp

Du Bois, J. W. (2007). The Stance Triangle. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 139–182). Benjamins. https://doi.org/10.1075/pbns.164.07du

Duijn, M. van, & Verhagen, A. (2019). Recursive embedding of viewpoints, irregularity, and the role for a flexible framework. *Pragmatics*, *29*(2), 198–225. https://doi.org/10.1075/prag.18049.van

Eisterhold, J., Attardo, S., & Boxer, D. (2006). Reactions to irony in discourse: Evidence for the least disruption principle. *Journal of Pragmatics*, *38*(8), 1239–1256. https://doi.org/10.1016/j.pragma.2004.12.003

Geeraerts, D. (2021). Second-order empathy and pragmatic ambiguity. In A. Soares da Silva (Ed.), *Figurative Language. Intersubjectivity and Usage*. John Benjamis, 19–40.

Gibbs, R. W. (2000). Irony in Talk Among Friends. *Metaphor and Symbol*, *15*(1), 5–27. https://doi.org/10.1080/10926488.2000.9678862

Giora, R. (1995). On irony and negation. *Discourse Processes*, *19*(2), 239–264. https://doi.org/10.1080/01638539509544916

Gironzetti, E., Attardo, S., & Pickering, L. (2016). Smiling, gaze, and humor in conversation. In L. Ruiz-Gurillo (Ed.), *Metapragmatics of Humor: Current research trends*. John Benjamins Publishing. https://doi.org/10.1075/ivitra.14.12gir

González-Fuente, S., Escandell-Vidal, V., & Prieto, P. (2015). Gestural codas pave the way to the understanding of verbal irony. *Journal of Pragmatics*, *90*, 26–47. https://doi.org/10.1016/j.pragma.2015.10.002

Goodwin, M. H., & Goodwin, C. (1986). *Gesture and coparticipation in the activity of searching for a word*. *62*(1–2), 51–76. https://doi.org/10.1515/semi.1986.62.1-2.51

Gullberg, M., & Kita, S. (2009). Attention to Speech-Accompanying Gestures: Eye Movements and Information Uptake. *Journal of Nonverbal Behavior*, *33*(4), 251–277. https://doi.org/10.1007/s10919-009-0073-2

Holler, J., & Kendrick, K. H. (2015). Unaddressed participants' gaze in multi-person interaction: Optimizing recipiency. *Frontiers in Psychology*, *6*, 1–14. https://doi.org/10.3389/fpsyg.2015.00098

Holt, E. (2016). Laughter at Last: Playfulness and laughter in interaction. *Journal of Pragmatics*, *100*, 89–102. https://doi.org/10.1016/j.pragma.2016.04.012

Janzen, T. (2019). Shared spaces, shared mind: Connecting past and present viewpoints in American Sign Language narratives. *Cognitive Linguistics*, *30*(2), 253–279. https://doi.org/10.1515/cog-2018-0045

Jehoul, A., Brône, G., & Feyaerts, K. (2017). The shrug as marker of obviousness. *Linguistics Vanguard: Multimodal Online Journal*, *3*(1), 1–9. https://doi.org/10.1515/lingvan-2016-0082

Katz, A. (2017). The standard experimental approach to the study of irony: Let us not be hasty in throwing out the baby with the bathwater. In A. Athanasiadou & H. L. Colston (Eds.), *Irony in Language Use and Communication* (pp. 237–254). John Benjamins Publishing Company. https://doi.org/10.1075/ftl.1.12kat

Kihara, Y. (2005). The mental space structure of verbal irony. *Cognitive Linguistics*, *16*(3), 513–530. https://doi.org/10.1515/cogl.2005.16.3.513

Lopez-Ozieblo, R. (2020). Proposing a revised functional classification of pragmatic gestures. *Lingua*, *247*, 1–18. https://doi.org/10.1016/j.lingua.2020.102870

McIntyre, D. (2006). *Point of View in Plays: A Cognitive Stylistic Approach to Viewpoint in Drama and Other Text-types*. John Benjamins Publishing. https://doi.org/10.1075/lal.3

Müller, C., & Posner, R. (2004). *The semantics and pragmatics of everyday gestures: Proceedings of the Berlin conference April 1998*. Weidler.

Panzeri, F., Giustolisi, B., & Mantovan, L. (2019). Joker face. Recognizing irony in the visual mode in spoken and signed languages. In J. Schloeder, D. McHugh, & F. Roelofsen (Eds.), *Proceedings of the 22nd Amsterdam Colloquium* (pp. 603–611).

Parrill, F. (2012). Interactions between discourse status and viewpoint in co-speech gesture. In B. Dancygier & E. Sweetser (Eds.), *Viewpoint in language: A multimodal perspective* (pp. 97–112). Cambridge University Press. https://doi.org/10.1017/CBO9781139084727.008

Partington, A. (2007). Irony and reversal of evaluation. *Journal of Pragmatics*, *39*(9), 1547–1569. https://doi.org/10.1016/j.pragma.2007.04.009

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U., Schütte, W., & Uhmann, S. (2009). Gesprächsanalytisches Transkriptionssystem (GAT 2). *Gesprächsforsch*, *10*, 152–183.

Sidnell, J. (2006). Coordinating Gesture, Talk, and Gaze in Reenactments. *Research on Language and Social Interaction*, *39*(4), 377–409. https://doi.org/10.1207/s15327973rlsi3904_2

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

Streeck, J. (1994). Gesture as Communication II: The Audience as Co-Author. *Research on Language & Social Interaction*, *27*(3), 239–267. https://doi.org/10.1207/s15327973rlsi2703_5

Sweetser, E. (2012). Introduction: Viewpoint and perspective in language and gesture, from the Ground down. In B. Dancygier & E. Sweetser (Eds.), *Viewpoint in Language: A Multimodal Perspective* (pp. 1–22). Cambridge University Press. https://doi.org/10.1017/CBO9781139084727.002

Sweetser, E., & Stec, K. (2016). Maintaining multiple viewpoints with gaze. In B. Dancygier, W. Lu, & A. Verhagen (Eds.), *Viewpoint and the fabric of meaning: Form and use of viewpoint tools across languages and modalities* (pp. 237–257). De Gruyter Mouton. https://doi.org/10.1515/9783110365467-011

Tabacaru, S. (2019). *A Multimodal Study of Sarcasm in Interactional Humor*. De Gruyter. https://doi.org/10.1515/9783110629446

Tabacaru, S., & Lemmens, M. (2014). Raised eyebrows as gestural triggers in humour: The case of sarcasm and hyper-understanding.

The European Journal of Humour Research, *2*(2), 11–31. https://doi.org/10.7592/EJHR2014.2.2.tabacaru

Thompson, S. A., & Suzuki, R. (2014). Reenactments in conversation: Gaze and recipiency. *Discourse Studies*, *16*(6), 816–846. https://doi.org/10.1177/1461445614546259

Tobin, V. (2018). *Elements of Surprise: Our Mental Limits and the Satisfactions of Plot*. Harvard University Press. https://doi.org/10.4159/9780674919570

Tobin, V., & Israel, M. (2012). Irony as a viewpoint phenomenon. In B. Dancygier & E. Sweetser (Eds.), *Viewpoint in Language* (pp. 25–46). Cambridge University Press. https://doi.org/10.1017/CBO9781139084727.004

Tomasello, M. (2008). *Origins of Human Communication*. MIT Press. https://doi.org/10.7551/mitpress/7551.001.0001

van Krieken, K., Sanders, J., & Sweetser, E. (2019). Linguistic and cognitive representation of time and viewpoint in narrative discourse. *Cognitive Linguistics*, *30*(2), 243–251. https://doi.org/10.1515/cog-2018-0107

Vandelanotte, L. (2017). Viewpoint. In B. Dancygier (Ed.), *The Cambridge Handbook of Cognitive Linguistics* (pp. 157–171). Cambridge University Press. https://doi.org/10.1017/9781316339732.011

Vandelanotte, L. (2019). Changing perspectives: Something old, something new. *Pragmatics*, *29*(2), 170–197. https://doi.org/10.1075/prag.18046.van

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. https://doi.org/10.2466/pms.108.2.565-572

Wittenburg, P., Brugman, H., Russel, A., Klassman, A., & Sloetjes, H. (2006). *ELAN: a Professional Framework for Multimodality Research*. Fifth International Conference on Language Resources and Evaluation., Genoa, Italy.

Zeman, S. (2019). The emergence of viewpoints in multiple perspective constructions. *Pragmatics*, *29*(2), 226–249. https://doi.org/10.1075/prag.18051.zem