As if that wasn't enough: English *as if* clauses as multimodal utterance constructions

CLAUDIA LEHMANN [ID]
University of Bremen

The following article reports on a multimodal corpus study of English *as if* constructions. The results of this study suggest that formulaic and insubordinate *as if* constructions are prosodically chunked as clauses, with formulaic *as if* constructions uttered with significantly higher pitch and insubordinate *as if* constructions with lower pitch when being compared with subordinate uses. In addition, insubordinate *as if* clauses are occasionally accompanied by frowns. It is argued that, although both constructions convey an ironic interpretation, multimodal markers of irony play only a minor role in explaining the findings. Instead, it is argued that the non-verbal features are construction-specific and can reasonably be explained as cross-modal collostructions. As such, the present article provides a description of the non-verbal features accompanying English *as if* clauses and provides a theoretical explanation. In doing so, some modest evidence for a multimodal Utterance Construction Grammar is also presented.

Keywords: insubordination, crossmodal collostructions, Utterance Construction Grammar, stance-related constructions, (Multimodal) Construction Grammar

1 Introduction

Construction Grammar assumes that constructions, i.e. 'learned pairings of form with semantic or discourse function' (Goldberg 2006: 5), are the basic units of language-related knowledge. The term *form* here implicitly refers to a *verbal form* since Goldberg continues by explaining that the definition of construction includes 'morphemes or words, idioms, partially lexically filled and fully general phrasal patterns' (Goldberg 2006: 5). However, face-to-face communication is inherently multimodal and some research in the past decade has questioned this bias toward monomodal verbal constructions and speaks up for a multimodal orientation in Construction Grammar (see e.g. Steen & Turner 2013; Zima & Bergs 2017).

The fact that spoken interactions are inherently multimodal is well received nowadays. Early studies on gestures and their relation to language showed that speech and gesture are closely time-aligned and that both contribute meaning to the utterance, i.e. are not redundant (Kendon 2004; McNeill 2005). Others soon followed this lead and, today,

there is a considerable body of research on how gestures contribute to utterance meanings (comprehensive reviews can be found in Vigliocco, Perniss & Vinson 2014; Feyaerts, Brône & Oben 2017; Perniss 2018). Gestures, it seems, are well-established semiotic resources that engage with semantic and pragmatic meanings. Recent advances in linguistically informed, multimodal studies suggest that the same can be claimed for semiotic resources other than (manual) gestures as well (see e.g. Feyaerts *et al.* 2022 on facial expressions as response turns).

In contrast to this, the relation between grammar and semiotic resources other than language is a controversial issue. In particular, among Construction Grammarians, the notion of MULTIMODAL CONSTRUCTION, i.e. a form–meaning pairing whose formal features comprise more non-verbal aspects, is disputed even though there are quite a few studies suggesting a close relation between syntactic, prosodic and kinesic properties for some constructions: Elvira-García (2019) shows how the intonation contour disambiguates elliptical and independent Spanish *si* + indicative clauses; Zima (2017) shows frequent co-occurrences of distinct gestures with semantic aspects of the [all the way from X PREP Y] construction; and Hinnell (2018) shows a distinctive and iconic relation between the use of manual gestures and aspect-marking constructions, to name but a few. Independently, Ward (2019) introduces the notion of PROSODIC CONSTRUCTION, i.e. 'a temporal configuration of prosodic features' with a meaning that is 'not necessarily closely aligned with words' (Ward 2019: 24). In doing so, he extends the notion of construction to prosodic forms like the CONSIDER THIS construction, which is characterized by a prosodically highlighted beginning (high pitch, loud, slow), then followed by narrow pitch range and, finally, ends in high pitch, which is typically used to provide further information to the hearer for them to consider in an argument (Ward 2019: 5–24).

A particular type of multimodal construction that seems to play a vital role for the present purposes is stance-related multimodal constructions. Stance is 'a public act by a social actor, achieved dialogically through overt communicative means, of simultaneously evaluating objects, positioning subjects (self and others), and aligning with other subjects, with respect to any salient dimension of the sociocultural field' (Du Bois 2007: 163) and, thus, stance-related constructions are *communicative means* as mentioned in the definition. A stance-related construction that has received a lot of attention is the shrug. Shrugs show formal variation (shoulder shrug, mouth shrug, head tilts, raised eyebrows, etc.) and can be used to indicate a lack of knowledge, obviousness, or disengagement (Streeck 2009; Debras & Cienki 2012; Debras 2017, 2021; Jehoul, Brône & Feyaerts 2017). Other stance-related constructions include the NEGATIVE ASSESSMENT CONSTRUCTION (Bressem & Müller 2017), which is instantiated by a throwaway gesture, and discourse management gestures that indicate strong disagreement with the interlocutor, such as pushbacks or pointing gestures that invade the shared discourse space (Wehling 2017).

While the studies reviewed above show a frequent co-occurrence of verbal and non-verbal resources, none of these co-occurrences seems to be obligatory in the strict sense. This has led some Construction Grammarians to argue that a statistically

sufficient frequency of co-occurrence does not mean that multimodal constructions are a cognitive reality and that possible candidates for multimodal constructions need to survive a deletion test to pass for genuine multimodal constructions (see Ningelgen & Auer 2017; Ziem 2017). Similarly, Hoffmann (2017) argues that, in most cases, constructions are unimodal but can be combined on-line while speaking, resulting in multimodal instantiations (see also Goodwin 2017 for his notion of contextual configurations). This view resonates partially with Uhrig's (2018) notion of CROSSMODAL COLLOSTRUCTIONS, i.e. strong associations between semiotically different constructions. Essentially, crossmodal collostructions require the independent existence of non-verbal form–function pairings, which can be combined with verbal form–function pairings. Examples of such independent, non-verbal form–function pairings include the shrug and the throwaway gesture, but also prosodic constructions as reviewed above.

Independent of the question that multimodal instantiations pose, some constructional approaches favor a prototype approach to constructions in general (see e.g. Gries 2003; Imo 2007; Schoonjans 2018). In this spirit, Cienki (2017) develops a prototype approach to multimodal constructions. He proposes that utterance constructions lie at the heart of spoken language analysis. These utterance constructions have a deep structure, which is stored in the constructicon and which contains information on 'tools that can be drawn upon to express the construction' (Cienki 2017: 3), i.e. the surface structure. These pieces of information can be verbal or non-verbal and can be more strongly or weakly associated with the utterance construction. The surface structure of the utterance construction is then a selection of relevant verbal and/or non-verbal behaviors and therefore, more often than not, stands in a metonymical relation to its deep structure. These relevant non-verbal behaviors associated with the utterance construction may also be conceptualized as crossmodal collostructions provided that there is an independently existing non-verbal construction. However, if the construction in question is associated with a non-verbal feature that does not have any independent meaning, this feature must be an integral part of the utterance construction even though it might not surface in every instance of it. In other words, if such a feature can be found, it supports the notion of multimodal constructions. Given the still thin empirical grounds, most researchers in Multimodal Construction Grammar agree on the fact that more empirical work needs to be done to come to any verifiable conclusions on the status of multimodal constructions (Hoffmann 2017; Schoonjans 2017).

The objective of the present article is twofold. The first is to provide empirical data by exploring English *as if* clauses and the non-verbal features they are frequently accompanied by. The second is to show that all of these features can be explained by resorting to the notion of crossmodal collostructions. Yet it will be argued that the assumptions of Utterance Construction Grammar are useful assets in explaining the different predictive power of these features. The view taken in this article is that both Utterance Construction Grammar and the notion of crossmodal collostructions are not mutually exclusive but complement each other in significant ways. Section 2 will show that *as if* constructions provide valuable insights for the discussion on multimodal

constructions, since non-verbal features seem to be necessary for their disambiguation in at least some cases. In section 3, previous research on multimodal markers relevant for the discussion of *as if* clauses will be reviewed. Section 4 provides the details on the quantitative multimodal corpus study that was conducted to gather empirical evidence for non-verbal features frequently accompanying *as if* clauses. Section 5 presents the results. Section 6 discusses the relation between the constructions' communicative function and the multimodal features they are accompanied by. And, finally, section 7 draws some conclusions for (Multimodal) Construction Grammar.

2 *As if* clauses

The use of English *as if* is a case in point to illustrate various degrees of what has been called INSUBORDINATION, i.e. 'the conventionalized main clause use of what, on prima facie grounds, appear to be formally subordinate clauses' (Evans 2007: 367). Examples (1) to (5) below illustrate the attested uses of English *as if* clauses, retrieved from the NewsScape Library of International Television News (Steen & Turner 2013).[1] Details on this archive and on the collection procedure will be provided in section 4.1. The video files from which the examples have been taken are available on the OSF platform (https://osf.io/usgw4/files/). Examples (1) and (2) illustrate genuine uses of *as if* clauses as subordinate clauses, while examples (3) to (5) illustrate uses with syntactic independence:

(1) It allowed him to move around as if this was Clarence Darrow in
     the courtroom (NewsScape
     2019-01-25_0300_US_MSNBC_The_Last_Word_With_Lawrence_
     ODonnell, 0:02:09-0:02:19; click to view or scan QR code)

(2) Justice Ginsburg passed away less than 48 hours ago, but it seems
     as if this is moving very fast and we could have a nominee very soon
     (NewsScape 2020-09-20_1500_US_KNBC_Meet_the_
     Press, 0:03:39-0:03:54, click to view or scan QR code)

(3) I should have known you would use the video of Cuomo coming up
     from his basement. As if that wasn't a propaganda video
     (NewsScape 2020-10-06_2100_US_FOX-News_The_Five,
     0:17:07-0:17:17; click to view or scan QR code)

---

[1] Examples (1) to (5) have been transcribed orthographically, using standard punctuation. They were simplified here to highlight their syntactic properties.

(4) As if this year hasn't been enough (NewsScape 2020-07-31_1800_
    US_KCBS_CBS _2_News_at_11AM, 0:17:49-0:17:53;
    click to view or scan QR code)

(5) He thought delaying me would make Republicans like me better.
    Yeah, right. As if (NewsScape 2010-03-09_0200_US_MSNBC_The_
    Rachel_Maddow_Show, 0:58:47-0:58:58; click to view or
    scan QR code)

Example (1) and (2) illustrate the use of *as if* as a subordinating conjunction. In example (1) *as if* introduces an adjunct adverbial clause. It is attached to a (syntactically independent) matrix clause and functions as a manner adjunct. In example (2), the *as if* clause functions as a complement to the verb *seems* and, semantically, it introduces a possibility of 'medium strength epistemic modality' (Huddleston & Pullum *et al.* 2002: 1152). In addition to these two uses, Huddleston & Pullum *et al.* (2002) also list two further functions, manner complement and adjunct of comparison (not illustrated here). Crucially, in all cases, the *as if* clause alone in this context would be ungrammatical without the matrix clause. Moreover, it introduces a proposition the speaker finds likely, but does not fully commit to. Given these syntactic and semantic commonalities, these uses will be referred to cumulatively as *subordinate* uses in the remainder of this article. Even though finer-grained analyses are possible here, treating these cases as one construction was considered feasible for the present purposes.

Example (3) is a prime example of insubordination. Here, the *as if* clause is syntactically independent. The speaker does not mean to say that *you would use that video as if that wasn't a propaganda video*, but she issues an afterthought and thereby expresses a negative attitude toward Cuomo's video. In other words, there is no matrix clause the *as if* clause could be attached to, and, nonetheless, it is grammatical. Like (2), it introduces a possibility, but this possibility is presented as pretty unlikely (from the speaker's perspective), if not even counterfactual. Pragmatically, the rejection of the state-of-affairs presented in the content clause usually receives an ironic interpretation (Brinton 2014: 96). In pragmatic terms, irony is an attributive language use, i.e. the speaker puts forward a proposition that alludes to an utterance or belief of some other person or some other version of themselves. This kind of attributive language use differs from other kinds in that it simultaneously expresses a dissociative attitude toward the proposition presented (see Wilson & Sperber 2012). Example (3) fits this definition of irony: the speaker attributes the thought that *that wasn't a propaganda video* to some other people (referred to as *you* in the previous clause and, presumably, the staff of politician Andrew Cuomo, who produced the video) and, at the same time, expresses a negative stance toward this thought, because the video represents a negatively connoted propaganda video in her view. Given the semantic resemblance between (2) and (3), Brinton (2014) and López-Couso & Méndez-Naya (2012) argue that the latter historically derives from the former.

Example (4) illustrates a special case of insubordinate *as if*. Like example (3), it is insubordinate, but it is lexically more constrained with the negated copula and the adjective *enough* being obligatory elements. In the NewsScape Library, between January 2018 and December 2020,[2] a total of 255 syntactically independent uses of *as if* clauses could be found, of which 74 (appr. 29%) showed this formulaic use. An independent treatment of such cases therefore seems feasible. What is more, meaning-wise, this use also behaves slightly differently from (3). While in (3) the content clause *that wasn't a propaganda video* is presented as unlikely, in example (4) *this year hasn't been enough* is presented as true (from the speaker's perspective), because further unfortunate events occurred, which are inferable from the context. Furthermore, this use of *as if* clauses often links two pieces of information, often bad news, while one is presented as 'the tip of the iceberg'. In example (4), it is the corona virus pandemic (*this year*) that is the backgrounded bad news while the hurricane, which is reported on, is presented as foregrounded bad news. In doing so, the speaker alludes to wishful thinking that everything is going to be fine while simultaneously expressing a negative stance toward the events reported, i.e. their frustration that the wishful thinking turned out to be wrong. From such a vantage point, these uses of *as if* clauses can be regarded as ironic, too, while the target of the ironic criticism is a different one: in (3), particular people are addressed (Cuomo's staff), while in (4) general expectations or beliefs are addressed. There are, however, instances of this formula, e.g. *as if this isn't exciting enough*, which express a playful, maybe slightly mocking rather than a negative attitude. In such cases, the ironic effect is less evident. In the following, this use will be labeled 'formulaic *as if*'.

And, finally, example (5) illustrates the use of *as if* as a bare complementizer. Here, *as if* is used independently in all respects – neither is it licensed by any element nor does it license any further elements. In contrast to the other two cases of insubordination, this use has found its way into the *Oxford English Dictionary* (2020), which states that it is 'Typically used as a sardonic response to a stated or reported suggestion' (s.v. *as*, adv. and conj.). This adequately describes example (5). *As if*, in combination with the ironic rejection *yeah, right*, assesses the thought reported previously in a negative way.

The expository paragraphs above are summarized in table 1.

The selection of examples (1) to (5) suggests that syntactic, semantic and pragmatic evidence can sufficiently disambiguate the different uses of *as if*. However, there are quite a few ambiguous cases (eleven in total), like the following:

(6) I just saw a video from Florida's beaches and they are absolutely
     packed with people sitting side by side and playing in the ocean as
     if this is not going on (NewsScape 2020-03-16_2100_US_
     MSNBC_MTP_Daily, 0:09:40-0:09:56; click to view or
     scan QR code)

---

[2] See section 4.1 for more details.

Table 1. *Overview on* as if *constructions from the NewsScape Library (January 2018 – December 2020) with token frequency*

Example(s)	Constructional label	Form	Function	Frequency
(1), (2)	subordinate *as if*	C *as if* C	introduce a proposition the speaker does not fully commit to	415
(3)	insubordinate *as if*	*As if* C	introduce a proposition the speaker rejects	181
(4)	formulaic *as if*	*As if* X BE not Y enough	introduce a proposition to which the speaker commits, but takes a distancing stance	74
(5)	bare *as if*	*As if*	reject a proposition put forward in response to some previous proposition	–[3]

Technically, example (6) could be an instance of a subordinate *as if* clause that introduces a counterfactual possibility, *with people sitting side by side and playing in the ocean* being the matrix clause(s). However, it could also be syntactically independent, semantically issuing an afterthought. Contrary to what could be expected, Lehmann & Bergs (2021) show that subordinate and insubordinate *as if* clauses don't show any difference as regards the tense of the verb. Essentially, the question of syntactic (in)dependence boils down to the question of which kind of attitude is conveyed in cases like (6). The obvious counterfactuality presented in the *as if* clause could either be treated as a neutral observation or a criticism of the way these people act, lending support to the analysis of (6) as subordinate or insubordinate, respectively. Since the attitude of the speaker is not indicated on the lexical level here, the disambiguation of (6) cannot be made on language-internal grounds, but relies on other modes than the verbal one.

In spoken language, this matter is complex and deserves systematic empirical attention – an objective that this article set itself to achieve. More specifically, this article explores the non-verbal resources speakers use in spoken interactions to mark *as if* constructions. The second objective, as already mentioned, is to square these empirical findings with the notion of multimodal constructions. There are quite a few possible outcomes:

1. Ironic *as if* clauses, irrespective of constructional type, could be associated with a particular set of non-verbal features when compared with non-ironic *as if* clauses. In such a case, there could be an independent set of non-verbal constructions that signals irony irrespective of verbal form.
2. Insubordinate and formulaic uses of *as if* clauses could be both associated with the same set of features when compared with subordinate uses. Such an outcome would

---

[3] Bare *as if* has been excluded from further considerations. The reasons for this will be provided in section 4.1.

lend support to the idea that there is an independent, non-verbal construction functioning as a marker for syntactic independence.

3. Insubordinate and formulaic uses of *as if* clauses could be associated with different sets of non-verbal features when compared with subordinate uses. Such an outcome supports an individual treatment of the two uses as different constructions. Moreover, depending on the exact nature of these non-verbal features, further conclusions might be drawn:

   a. If these non-verbal features have been described elsewhere, fulfilling the same or a similar function, it could be assumed that these non-verbal features present an independent non-verbal construction that is associated with the verbal construction (i.e. forms a crossmodal collostruction).

   b. If there are non-verbal features that have not been attested elsewhere (and are unlikely to function in a similar way independent of the particular verbal construction), these provide modest evidence for a genuine multimodal construction.

3 Multimodal aspects of subordination and irony

As was shown above, *as if* constructions vary largely on two grounds, i.e. as regards their constructional complexity and the stance the speaker takes toward the proposition expressed in the *as if* clause. Therefore, a review of the literature on multimodal markers of (in)subordination and irony will be provided, which served to delimit the variables for the empirical study outlined in section 4.

3.1 *Marking of dependent and independent clauses*

Early findings on the prosody–syntax interface observed that dependent structures are often realized with a rising pitch contour, which signals that more is going to follow (Bolinger 1984; Wells 2006). More recent, empirical studies confirm this observation (Lelandais & Ferré 2016, 2017, 2019; Elvira-García, Roseano & Fernández-Planas 2017; Elvira-García 2019; Maschler 2020). Moreover, these studies show that dependent structures can also be intonationally integrated into their host structure (Lelandais & Ferré 2016), are usually slower than more independent structures and tend to be accompanied by silent pauses (Lelandais & Ferré 2016, 2019; Köhn, Baumann & Dörfler 2018). Other findings are less consistent. While Köhn, Baumann & Dörfler (2018) report a lower mean pitch for German subordinate clauses, Lelandais & Ferré (2016) find a lowered mean pitch for English appositive clauses. Furthermore, Köhn, Baumann & Dörfler (2018) report a lowered intensity[4] for subordinate clauses, while Elvira-García (2019) finds no effect of intensity on the discrimination of elliptical and independent Spanish *sí*-clauses.

---

[4] Intensity is the acoustic correlate of loudness.

Empirical studies concerned with kinesic information accompanying syntactic (in)dependence are rare. Notable exceptions are Lelandais & Ferré (2017, 2019). They report that syntactically independent structures are kinesically set off from their surrounding co-text, i.e. they are often produced with non-overlapping, distinct manual gestures, gaze changes and eyebrow rises. Dependent structures, on the other hand, are often produced with overlapping hand gestures, thus creating a kinesic link to their hosts (see also Maschler 2020 for similar observations).

3.2   Marking of irony

Research on the so-called 'ironic tone of voice' has not been conclusive to date. One reason for the controversial findings is that the prosodic marking of irony is language-specific (on the difference between the prosodic profiles for English and Cantonese irony see Cheang & Pell 2009, 2011). Another reason is that scripted and unscripted irony seems to trigger different prosodic profiles (Rockwell 2000). Research on unscripted English suggests that it is marked by a slower tempo (Rockwell 2007; Bryant 2010), a higher mean pitch (Bryant & Fox Tree 2005; Rockwell 2007), greater pitch variability (Rockwell 2007) and greater intensity variability (Bryant & Fox Tree 2005).

Kinesic cues to an ironic meaning are similarly controversial. Colston (2020) maintains that gaze aversion is a feature of irony, while Caucci & Kreuz (2012) report looks to the recipient to accompany an ironic remark. Other kinesic features that are sometimes reported to accompany irony include raised eyebrows and frowns (Tabacaru & Lemmens 2014; Tabacaru 2019, 2020), rapid blinking (Kreuz 2020), tightened lips, smiles and laughter (Caucci & Kreuz 2012) as well as head nods (Caucci & Kreuz 2012; Tabacaru 2019) and head tilts (Tabacaru 2019). In contrast to these findings, Attardo *et al.* (2003) find the so-called 'blank face' to be prominent in their data.

4   Study details: method and annotations

Given the fact that insubordinate *as if* clauses are syntactically independent structures conveying an ironic meaning, the review above suggests that they are most likely realized in a separate tone-unit and accompanied by eyebrow rises. Apart from these, no further commonalities are noted in the literature between multimodal markers of independent syntactic structures and irony. Interestingly, there are also some cues in conflict here, i.e. speech tempo and overall mean pitch: while ironic utterances are usually slower and lower in pitch, independent syntactic structures tend to be faster and higher in pitch. In the following, the details of a corpus study that investigates if and how subordinate and insubordinate *as if* clauses are multimodally marked in naturally occurring interactions are laid out.

The study is a corpus-based analysis of multimodal markers of irony. The multimodal archive used here is the UCLA NewsScape Library of International Television News (Steen & Turner 2013). This archive contains a collection of digitized television news

programs. The collection extends from 2004 and runs until the present day. In March 2021, it already counted 409,532 hours of programming of American English television containing 2.94 billion words and is updated on a daily basis (Uhrig 2021). The video files provided by the NewsScape Library include useful information on prosody, facial expressions, head movements and manual gestures. The archive was accessed through the facilities of the Distributed Little Red Hen Lab (which is co-directed by Francis Steen and Mark Turner), using the Edge search engine.

    Given this huge archive, the search was limited to video files from January 2018 to December 2020 and to the string *as if this/that*. Using *as if* as the only search terms resulted in a considerable imbalance toward subordinate constructions. Lehmann & Bergs (2021) suggest that subordinate *as if* clauses are associated with proximal demonstrative pronouns, while insubordinate *as if* clauses are associated with distal demonstrative pronouns. Therefore, including these kinds of pronouns in the search string was considered a useful limitation of the data. The results obtained in this way were further limited. Videos in which there was a considerable amount of overlapping speech or noise or in which the speaker's face was not visible were excluded from further analyses. Ambiguous cases that could not be assigned to a construction on syntactic grounds (like example (6) above) were not included in the analyses either. This procedure resulted in a total of 668 hits.

    These hits were annotated for interaction type, construction, speaker identity, interpretation, and the syntactic form of the prosodic chunk in which the *as if* was embedded. Interaction type was categorized as either 'scripted' (TV series, movies, commercials), 'monologue' (stand-up routines, news reports), 'video call' or 'face-to-face interaction' since previous studies reported different markers of irony for scripted and unscripted types of interactions. As for construction, values were 'subordinate', 'insubordinate' and 'formulaic'. The bare *as if* construction had to be excluded from the analysis due to the methodological considerations described above, which included using the search string *as if this/that*. Speaker identity was annotated manually due to the fact that the name of the speaker is not always provided in the metadata files of the NewsScape archive and needed to be extracted from the text included in the video files. If no information on the identity of the speaker was provided in the video (e.g. in the case of street interviews), the speakers were labeled as anonymous and numbered consecutively. Interpretation was categorized as 'ironic' or 'non-ironic' based on the definition by Wilson & Sperber (2012). Examples ambiguous between an ironic and non-ironic interpretation were excluded. Prosodic chunks are difficult to identify because they are fuzzy entities (Barth-Weingarten 2016). Thus, boundaries between prosodic chunks were determined using a variety of features, including pauses and inbreaths (Szczepek Reed 2011), falling pitch, lengthening and voice creaks (Barth-Weingarten 2016). The chunks containing *as if* were then annotated for their syntactic form, the values being 'sentence', 'clause', 'verb phrase', 'as if' and 'other'.

    In order to identify prosodic features, the video files were converted to wav format and analyzed with Praat (Boersma & Weenink 2019). This software was used to measure

pauses before and after the prosodic chunk as well as internal pauses. Furthermore, duration per syllable, mean pitch, standard deviation of mean pitch as well as pitch range (maximum minus minimum pitch) of the prosodic chunk containing *as if* was measured. Although intensity (i.e. the acoustic correlate of loudness) would have been interesting to investigate, its measurement outside the laboratory is highly unreliable and, therefore, is not considered any further in this study. Moreover, a pilot analysis of the first 100 hits revealed that all of them were produced with non-rising intonation, which is why intonation contour is not further considered in this study. One reason for this observation may be an outcome of the search string used: the demonstrative pronouns *this* and *that* are, mostly, used anaphorically, i.e. subordinate *as if* clauses follow the matrix clause and, thus, non-rising intonation is more likely.

In addition to prosodic features, the data were also annotated for kinesic features by the author of this article. To do so, the videos were paused at the onset of the *as if* clause and then viewed frame by frame. The features under consideration were gaze direction, head movements, blinking rate as well as movements in the eye, eyebrow and mouth region based on a subset of action units described in Ekman & Friesen (2003). Gaze direction was annotated broadly as either 'directed at the camera', 'to recipient', 'to the audience', 'to an object', or 'elsewhere'. Measurements of gaze direction are imprecise without eye-tracking techniques, but here the perspective of the (uninitiated) viewer was taken. Head movements were categorized as 'nod', 'shake', 'tilt', 'turn', 'none' or 'other'. The blinking rate was determined by counting the number of blinks during the utterance of the prosodic chunk divided by its total duration. Movements in the eye region were categorized as 'blinking', 'closed', 'upper lid raised', 'lower lid raised', 'cheeks raised', 'other' and 'none'. Movements in the eyebrow region were categorized as 'raised', 'frowning', a 'combination' thereof, 'other' and 'none'. Finally, movements in the mouth region were categorized as 'smile or laughter', 'other' and 'none'. Another variable was 'blank face'. Since there is no agreed-upon definition of what counts as a blank face, this variable received a positive value when all facial action units received a 'none'-value. In all other cases, it received a negative value.

An overview on all annotated values can be found in table 2. The data are made available at https://osf.io/usgw4/files/

Even though the review presented in section 3 suggests that manual gestures might be relevant for discriminating between syntactically dependent and independent *as if* constructions, these were neglected in the present study for two main reasons. First of all, analyzing manual gestures requires the hands of the speaker to be visible to the researcher. This was the case for less than half of the data. What is more, however, a preliminary view of the data did not suggest that the speakers observed in the NewsScape Library gesture a lot when uttering *as if* clauses. The reasons for this can only be speculated about. One reason might be that most of the speakers are experienced TV personalities (often news anchors) who are probably aware of the fact that their hands might not be visible and who might receive some formal training in non-verbal communication. Irrespective of these reasons, manual gestures were excluded from further analyses in order to get a sufficiently large dataset for more promising features.

Table 2. *An overview on the variables and values used in the study*

Contextual variables

Variable	Values
Interaction type	Scripted, monologue, video call, face-to-face
Speaker	Ab Stoddard, Abby Phillip, …, Zerlina Maxwell

Textual variables

Variable	Values
Construction	Insubordinate, subordinate, formulaic
Syntactic form of prosodic chunk	Sentence, clause, verb phrase, as if, other
Interpretation	Ironic, non-ironic

Kinesic variables

Variable	Values/Measured
Gaze	To camera, to recipient, to audience, to object, elsewhere
Head movement	Nod, shake, tilt, turn, other
Blinking rate	Number of blinks/duration of prosodic chunk
Eyebrow region	Raised, frown, combination, other, none
Eye region	Blinking, closed, upper lid raised, lower lid raised, cheeks raised, other and none
Mouth region	Smile or laughter, other, none
Blank face	Yes, no

Prosodic variables

Variable	Measurement	Measured in
Pause before prosodic chunk	(Automatic extraction)	milliseconds (ms)
Pause after prosodic chunk	(Automatic extraction)	ms
Internal pause	(Automatic extraction)	ms
Duration per syllable	Total duration divided by the number of syllables of the prosodic chunk	ms
Mean pitch	(Automatic extraction)	Hertz (Hz)
Standard deviation from mean pitch	(Automatic extraction)	Hz
Pitch range	$Pitch_{max}$ minus $Pitch_{min}$	Hz

5 Results

In the following section, the results of the corpus study are presented. First, some observations are made on the constructions themselves. This is followed by a presentation of the results of linear mixed-effects models run using the lme4-package (Bates *et al.* 2015) for irony and the mclogit-package (Elff 2022) for *as if* constructions in the statistics program R (R Core Team 2019).
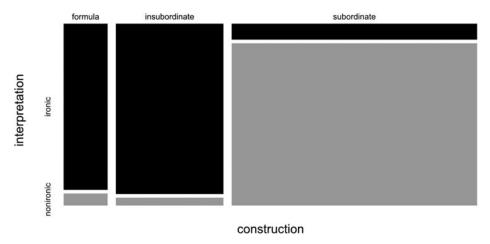
Figure 1. Mosaic plot illustrating the frequencies of *as if* constructions in spoken English and their interpretation

5.1   *As if* constructions and irony

The corpus data gained from the UCLA NewsScape Library of International Television News confirms that subordinate *as if* constructions typically convey a non-ironic meaning, while insubordinate constructions convey an ironic meaning. The formulaic *as if* construction is also more often than not used ironically, but less frequently than the insubordinate construction. This is illustrated in figure 1.

5.2   Modeling ironic *as if* clauses

Since independent *as if* clauses, formulaic or insubordinate, often convey an ironic meaning, one question is whether there is a prosodic or kinesic profile that is associated with irony irrespective of the construction used. Thus, a model using the lme4 package (Bates *et al.* 2015) in R was fitted that identified multimodal markers of irony conveyed by *as if* clauses, ignoring construction as a potential factor. The R script is made available on the OSF platform (https://osf.io/usgw4/files/). The results are summarized in table 3.

Table 3 shows that the interpretation of an *as if* clause as ironic and its multimodal marking show significant variance in intercepts across speakers and interaction types. In addition, ironic *as if* clauses are significantly more often prosodically chunked as clauses and are significantly faster than non-ironic *as if* clauses. None of the other features listed in the model (i.e. blinking rate, pausing, mean pitch, gaze behavior and movements in the eye region) reached a significant level, even though they improved the model fit. The features not listed in table 3 did not improve the model fit, i.e. they seem to have no influence on the interpretation of an *as if* clause as (non)ironic. The odds ratios and their confidence intervals are illustrated in figure 2.

Table 3. *Summary of the final model for irony*

MODEL INFO

Observations: 538
Dependent Variable: isironic
Type: Mixed effects generalized linear regression
Error Distribution: binomial
Link function: logit

MODEL FIT

AIC = 511.11, BIC = 562.57
Pseudo−R^2 (fixed effects) = 0.44
Pseudo−R^2 (total) = 0.60

FIXED EFFECTS

	Est.	S.E.	z val.	p
Intercept	−2.12	0.57	−3.74	0.00
Prosodic chunk = clause (isclause)	3.44	0.40	8.50	0.00
Blinking rate (blinking)	−0.05	0.19	−0.27	0.79
Pause after C (pause3)	0.22	0.16	1.36	0.17
Mean pitch (pitch)	−0.14	0.14	−1.00	0.32
Neutral eyes (neueyes)	−0.19	0.31	−0.60	0.55
Speech tempo (tempo)	−0.39	0.15	−2.66	0.01
Pause before C (pause1)	−0.14	0.21	−1.13	0.26
Internal pause (pause2)	−0.39	0.21	−1.86	0.06
Gaze to audience (audience)	1.99	1.80	1.11	0.27

RANDOM EFFECTS

Group	Parameter	Std. dev.
Speaker	(intercept)	0.91
Interaction type	(intercept)	0.64
Grouping variables:		
Group	#groups	ICC
Speaker	341	0.18
Interaction type	4	0.09

Figure 2 confirms that only prosodic chunking and speaking rate (tempo) can reliably predict ironic *as if* clauses since their confidence intervals do not cross the vertical zero-effect line. All of the other terms that have entered the model cannot confidently be used to predict an ironic interpretation, even though their odds ratios suggest some tendencies. These are the following: ironic *as if* clauses tend to be followed by a pause, but are not preceded by one, nor are there any internal pauses. Also, ironic *as if* clauses tend to be lower in pitch and tend to be accompanied by movements in the eye region, though not by frequent blinks. Finally, the speaker of an ironic *as if* clause tends to look more at the audience (if present) than when being non-ironic.
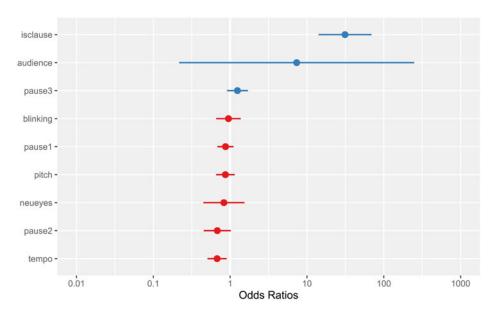
Figure 2. Odds ratios and confidence intervals for the model terms in descending order

5.3 *As if* clauses

To model *as if* constructions a polytomous model was fitted using the mclogit package (Elff 2022) with subordinate clauses as reference level. Table 4 summarizes the final model.

Table 4 shows that the relation between *as if* clauses and their multimodal markers is significantly influenced by the speaker and the interaction type. It also shows that prosodic chunking and mean pitch act as significant predictors of the different *as if* constructions. The relation between construction and prosodic chunking is further illustrated in figure 3.

Figure 3 illustrates that while subordinate constructions can be chunked prosodically in various ways, syntactically independent *as if* clauses (both insubordinate and formulaic) show a significant tendency to be chunked as clauses.

Despite this prosodic commonality, formulaic and insubordinate *as if* constructions differ in their mean pitches. That is, formulaic *as if* constructions are higher in pitch than subordinate constructions, while insubordinate constructions are lower in mean pitch than subordinate constructions. Other features that improved the model fit were the speaking rate (tempo) and frowning, but these did not reach a significant level. None of the other features improved the model fit. The estimates and their confidence intervals for the model are illustrated in figure 4.

Figure 4 illustrates the estimates and the confidence intervals for the model terms. It shows that prosodic chunking of *as if* clauses as clauses seems to be a reliable predictor for the formulaic and insubordinate constructions since their confidence

Table 4. *Summary of the final model for* as if *constructions*

MODEL INFO

Observations: 645
$N_{speaker.interaction}$: 462
Dependent Variable: construction
Type: Mixed effects generalized linear regression
Error Distribution: multinomial
Link function: logit

MODEL FIT

AIC = 198.16, BIC = 242.85

FORMULA VS SUBORDINATE

	Est.	S.E.	z val.	p
(Intercept)	−4.31	0.55	−7.78	7.32e−15
Frown	−0.86	0.65	−1.34	0.18
Pitch	0.49	0.21	2.36	0.02
Tempo	−0.38	0.23	1.64	0.10
Isclause	3.41	0.58	5.84	5.23e−09

INSUBORDINATE VS SUBORDINATE

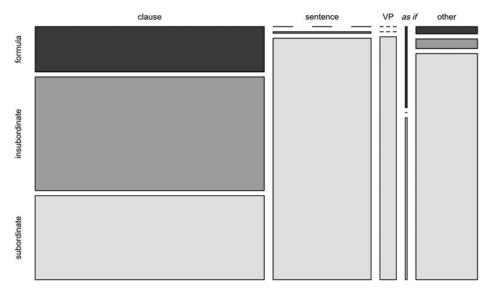|	Est.	S.E.	z val.	p
(Intercept)	−4.25	0.51	−8.32	<2e−16
Frown	0.55	0.43	1.27	0.20
Pitch	−0.46	0.17	2.72	0.01
Tempo	−0.26	0.16	1.56	0.12
Isclause	4.44	0.53	8.43	<2e−16

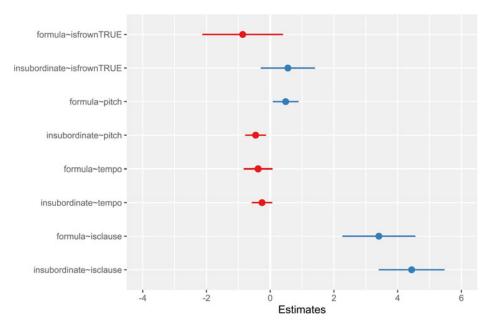Figure 3. Mosaic plot showing the relation between *as if* constructions and their prosodic chunking

Figure 4. Estimates and confidence intervals for the final model for *as if* constructions

Table 5. *Frequencies of the factor variables as well as means and standard deviations of the numeric variables in the final model for* as if *clauses*

	Subordinate		Insubordinate		Formulaic	
	Abs. freq.	Rel. freq.	Abs. freq.	Rel. freq.	Abs. freq.	Rel. freq.
Frown	62/413	15%	35/181	19%	6/74	8%
Clausal chunk	130/413	31%	176/181	97%	70/74	95%
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Pitch	153Hz	29	144Hz	30	164Hz	29
Tempo	218ms/ syllable	58	210ms/ syllable	49	190ms/ syllable	33

intervals do not cross (or even come near) the vertical zero-effect line. Both kinds of constructions tend to be chunked as clauses. As regards the mean pitches, figure 4 illustrates that formulaic *as if* tends to be uttered with a higher pitch, while insubordinate *as if* is most often uttered with lower pitch. As mentioned above, none of the other terms reached a significant level, but some tendencies can be observed.

Formulaic *as if*, for example, tends to be fast in tempo and is less likely to be accompanied by frowning, while insubordinate *as if* clauses also tend to be rather fast in tempo, but show a higher tendency to be accompanied by frowns. Table 5 provides some details on the absolute and relative frequencies of the factor variables (frowns, chunking) as well as mean and standard deviations of the numeric variables (mean pitch, tempo).

The results presented in this section will be illustrated and discussed in the following section.

6   Discussion

The most striking finding is that both irony and syntactically independent *as if* clauses are significantly associated with prosodic chunks that correspond to syntactic clauses, while non-ironic, subordinate *as if* clauses can be uttered in various ways. This variety of chunking of subordinate clauses is illustrated by the following examples:[5]

(7) uh- Justice Ginsburg is- is- is- uh passed away less than forty eight
     hours ago | but- it seems as if (.) uh THIS is moving very fast | and
     we could have a nominee VERY soon | what can you tell us
     (NewsScape 2020-09-20_1500_US_KNBC_ Meet_the_Press,
     0:03:39-0:03:54, click to view or scan QR code)[6]

(8) The president is campaigning | as if this pandemic is over | holding
     multiple rallies per day (NewsScape 2020-10-29_0900_
     US_CNN_Early_Start_With_ Christine_Romans_and_
     Laura_Jarrett, 0:04:53-0:4:59; click to view or scan QR code)

(9) But NOW it looks as if that tornado threat | is still going to be uh
     impactful across the deep south (NewsScape 2020-10-
     10_0900_US_CNN_CNN_Newsroom _Live, 0:12:52-0:13:02;
     click to view or scan QR code)

Example (7) illustrates a subordinate *as if* clause that is uttered together with the matrix clause as one prosodic chunk. With only two syllables, the matrix clause in this example is rather short and this might be the reason why further syllables are attached to the prosodic chunk. Prosodic integration of dependent structures into host structures has also been observed in Lelandais & Ferré (2016). In example (8), though, the matrix clause (*the president is campaigning*) consists of eight syllables and is thus considerably longer than the matrix clause in (7). This might be one reason why the speaker of (8) opted to chunk the *as if* clause in a separate prosodic unit. Another possible reason why the speaker has chunked the utterance like this becomes obvious when the video is consulted: when she utters the *as if* clause, the speaker is (slightly)

---

[5] Vertical bars indicate boundaries between prosodic chunks.
[6] Note that example (7) is the same as example (2), but repeated and renumbered here for the sake of convenience.

shaking her head. In doing so, she is not presenting the news in a neutral way, but indicates her stance toward the proposition expressed in the *as if* clause. Thus, chunking helps the speaker to provide the hearers with clues about the scope of the stance. Finally, example (9) illustrates a kind of chunking of subordinate *as if* clauses that is also quite common (N = 94). It has been categorized as 'other' in the present study, but is probably better described as a topic-comment structure with the topic being chunked as one prosodic unit and the comment in the other, following unit (see also Wells 2006: 72–3). In this example, the topic is the tornado threat, which is established in the first part of the utterance and is being commented on in the second prosodic chunk as still being impactful. From a syntactic point of view, in examples like this, the first prosodic chunk consists of the matrix clause, the *as if*, and the subject of the *as if* clause (usually in the form of a noun phrase), while the other prosodic chunk consists of the remaining elements of the *as if* clause. In contrast to example (7), the speaker of (9) can prosodically highlight both the subject and aspects of the remaining clause, while in (7) only one of the two (here: the subject) can be emphasized.

Syntactically independent (i.e. insubordinate and formulaic) *as if* clauses usually convey one proposition, which is also displayed as such prosodically, as the following examples illustrate:

(10) He just yells back | why you're the one always yelling the questions | as if that's something new (NewsScape 2020-11-21_0400_US_FOX-News_Fox_News_at_Night_With_Shannon_Bream, 0:37:19-0:37:25; click to view or scan QR code)

(11) And as if that wasn't enough | according to the Washington Post | the Ukrainians send a delegation to the White House in July (NewsScape 2019-10-31_0635_US_KABC_Jimmy_Kimmel_Live, 0:02:02-0:02:12; click to view or scan QR code)

In example (10), the speaker quotes some other person who presumably said *why you're the one always yelling the questions* and mocks this person by claiming that this is no new information. This criticism is presented in one prosodic chunk to sufficiently distinguish the quote from the speaker's own stance toward the quote. Likewise, in example (11), the formulaic *as if* construction is used to link two pieces of bad news and this link is prosodically set off from its surrounding material to emphasize that the following piece of information is just 'the tip of the iceberg' in a series of bad news.

These findings are in line with previous research on free constituents (Ford, Fox & Thompson 2002). Free constituents are syntactically and prosodically independent, but semantically related extensions of a previous utterance and are used to provide a stance toward it. Even though the free constituents described in Ford *et al.* (2002) are formally noun phrases, the observations made for them can be extended to syntactically independent *as if* clauses as well. Given these parallels, it could be argued that stance-related constructions tend to be chunked as one prosodic unit. In other

words, there could be an abstract construction with [prosodic chunk] on the formal side and [information package] on the meaning side and stance-related information being the particular kind of information conveyed here. Providing direct evidence for such an assumption lies outside the scope of the present article, but the examples above provide some indirect evidence supporting it. Essentially, then, one might assume a crossmodal collostruction between the prosodic chunk construction and syntactically independent *as if* constructions.

A feature that is significant for an ironic interpretation, but not for *as if* constructions is tempo. The direction of this finding is surprising because previous research suggested that syntactically dependent structures tend to be slower, but this could not be confirmed. Rather, ironic *as if* clauses tend to be faster than non-ironic ones. The statistical model reported above suggests that the speaker and the interaction type have an influence on the fixed effects of the model, increasing uncertainties. The interactional data used in Lelandais & Ferré (2016) are based on a limited number of participants and have been recorded in one setting. In contrast, the present study is based on 482 different combinations of speakers and interaction types. This use of different interaction types and speakers may explain why the findings could not be replicated. It is likewise surprising that ironic *as if* clauses, independent of the construction, are faster than non-ironic *as if* clauses, since previous research suggests that irony is slower than non-irony (see section 3.2 above). An alternative interpretation is provided in Ward (2019). Ward reports on a prosodic construction he calls INDIFFERENCE CONSTRUCTION (2019: 183–5), which is characterized by a fast tempo (among other things) and usually conveys the speaker's indifference toward their interlocutor's point of view. If this is the case, a prosodic construction is superimposed on a verbal construct, independent of the grammatical construction used, i.e. this is neither evidence for crossmodal collostructions nor multimodal constructions.

In any case, it is surprising that no other feature except prosodic chunking and tempo reached a significant level. Only two further features that entered the model correspond to the ones described in the literature, namely gaze aversion (here: looks to the audience rather than the addressee) and a (slightly lowered) pitch level, albeit non-significantly. None of the other features in the model have been reported to mark irony before or, if they have, not in the predicted direction. However, given the fact that previous research on multimodal markers of irony was also inconclusive or controversial, it might be possible that verbal irony is a heterogeneous phenomenon. More specifically, it seems that the ironic function is an umbrella term (see also Gibbs 2000; Simpson 2011) and that the function of supposedly ironic utterances needs more fine-grained analyses including precise descriptions of the stance conveyed. Essentially, there is no evidence of a set of multimodal features that are linked to an ironic interpretation (i.e. non-verbal 'irony' construction(s)) and that are, in turn, associated with *as if* constructions.

Apart from chunking, there are two further features that distinguish subordinate from insubordinate and formulaic *as if* constructions, respectively. These are mean pitch and frowning. According to the model, insubordinate constructions are comparatively low

in pitch and tend to be accompanied by frowns (albeit frowning was non-significant). The two features are illustrated by the following example:

(12) is now tweeting polls about |.hh America is losing faith in our
      democracy and our elections | as if this is winning for him | maybe
      it is (NewsScape 2020-11-19_0300_US_ CNN_CNN_Tonight_
      with_Don_Lemon, 0:15:00-0:15:13; click to view or scan QR code)

In this example, news anchor Don Lemon first quotes one of the former US president Donald Trump's tweets after he has lost the election to president-elect Joe Biden in November 2020. After quoting the tweet, Lemon comments on it by rejecting the idea that this might be winning for Donald Trump. However, in the next utterance, he then changes his mind and finds this idea more likely. Both the quote and the *as if* clause are accompanied by frowning, indicating that the speaker takes a negative stance toward these propositions, while, in the following utterance, his facial expressions become neutral. This use of frowning has already been observed elsewhere. Kaukomaa, Peräkylä & Ruusuvuori (2014), for example, show that turn-initial frowning foreshadows trouble talk including negative evaluations and disaffiliation. Given that such use of frowning has been observed elsewhere and given that, in the example, the speaker started frowning before uttering the *as if* clause, this is good evidence of a crossmodal collostruction.

The mean pitch in this example is also noticeably low. Figure 5 illustrates the pitch movements of the first and the second part of this example.

Figure 5 shows that the first part of the example, the quote, is rather high in pitch, with a mean pitch of 178 Hz. The second part, i.e. the *as if* clause, on the other hand, is low in comparison, with a mean pitch of 109 Hz. Since syntactically independent structures have been shown to be indicated by lower pitch (Lelandais & Ferré 2016), this might serve as one explanation for the lowered mean pitch of the insubordinate *as if* construction here. However, the construction is preceded by a pause of more than a second, and, thus,
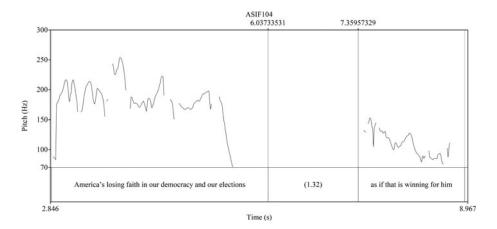

Figure 5. Pitch movements in example (12)

this explanation is unlikely. Traditionally, a lowered mean pitch has also been associated with the speaker's dominance and superiority due to the fact that tall people have longer larynxes and, therefore, lower voices (Gussenhoven 2004). This biologically motivated association still has an influence on how high and low voices are being perceived (see empirical findings in Puts *et al.* 2007). Therefore, assuming a prosodic construction with low pitch on the form side and dominance on the function side seems likely. In example (12), the speaker probably indicates his confidence when rejecting the idea that subverting the outcome of the elections is a winning strategy. To do so, he uses an insubordinate *as if* construction, which is matched with the 'low pitch construction'. Assuming a crossmodal collostruction in this case seems feasible.

Formulaic *as if* constructions, on the other hand, are accompanied by higher than normal pitch, which is illustrated in example (13).

(13) as if this wasn't enough news for today | the show is kind of:
        topsy-turvy | but let's talk about facebook
        (NewsScape 2018-04-09_2200_US_FOX-News_Special _
        Report_With_Bret_Baier, 0:51:31-0:51:37; click to view or
        scan QR code)

The pitch movements of this example are illustrated in figure 6.

Figure 6 shows that formulaic *as if* starts with a high onset (with a maximum of 241 Hz) and then gradually declines with a mean pitch of 154 Hz. The following two utterances are lower in mean pitch (with mean pitches of 131 Hz and 135 Hz, respectively). This finding is in contrast to previous findings on syntactically independent structures (see section 3.1 above). It might be argued that formulaic *as if* clauses in general, and example (13) in particular, are exceptional, because they occur at the beginning of the turn and, therefore, setting them off prosodically with a
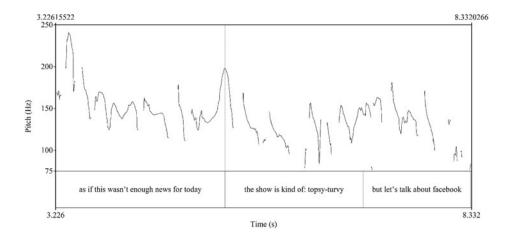

Figure 6. Pitch movements in example (13)

lowered pitch is unnecessary. Indeed, 27 percent of formulaic *as if* clauses (N=20) occur at turn beginnings. However, raises in pitch can also be observed for turn-internal formulaic *as if* constructions (see e.g. example (11) above). Rather than signaling turn beginnings, the speaker appeals to the audience when they use a higher voice. To be more precise, the prosodic aspects of formulaic *as if* fit what Ward (2019: 182) describes as the EMPATHY BID CONSTRUCTION, i.e. a configuration of prosodic features speakers use when telling a story that, from their perspective, deserves an empathetic uptake. The empathy bid construction is, among other things, characterized by raised pitch, articulated speech and increased loudness. While the latter two features cannot be measured reliably in non-laboratory settings, an informal perception of these confirms their presence in example (13). The function of the empathy bid construction is to seek empathy from the interlocutor. In example (13), being the news, there are no interlocutors, but it might be argued that the news anchor appeals to his audience's empathy. This finding suggests that speakers of formulaic *as if* constructions tend to bond with their interlocutors by seeking empathy rather than tending to distance themselves from the proposition expressed. Essentially, there is some evidence that formulaic *as if* constructions are associated with the prosodic empathy bid construction.

7   Summary and conclusion: multimodal *as if* constructions

In the closing paragraph of section 2, several possible outcomes of this study and their implications for a Multimodal Constructional Grammar were explored. The first possibility was that there might be a set of non-verbal forms indicating irony that match up with *as if* clauses and support their ironic interpretation, irrespective of the construction used. This possibility could not be confirmed. Even though this might have been expected, an ironic interpretation of the utterance does not have any explanatory power for the non-verbal features. There are only two features distinguishing ironic from non-ironic *as if* clauses in significant ways, i.e. prosodic chunking as a clause and a fast tempo, and these features alone cannot be considered sufficient for predicting ironic utterances. What is more, only one of the two features, prosodic chunking, was also significantly associated with *as if* constructions and this alone provides insufficient evidence for a non-verbal irony-construction matching up with *as if* constructions.

The second possible outcome stated that both kinds of syntactically independent *as if* constructions might be accompanied by the same non-verbal features due to the fact that they are both syntactically independent. However, the findings above have shown that there are subtle differences in the non-verbal markers that accompany syntactically independent *as if* clauses, even though both (insubordinate and formulaic *as if* clauses) are similar in verbal form and function: both constructions are syntactically independent and convey a distancing attitude toward some utterance or event mentioned in the previous context. Despite these formal and functional similarities, the two constructions differ significantly in their mean pitches. As a consequence, it seems that both formulaic and insubordinate *as if* constructions fulfill related, albeit

sufficiently different interactional functions and, crucially, seem to be accompanied by different non-verbal features.

The present study therefore supports the third possible outcome, i.e. that there are individual profiles for each *as if* construction. What is more, the study also supports the notion of crossmodal collostructions: all the features observed for insubordinate and formulaic *as if* constructions have been described to work in interaction with other constructions elsewhere, fulfilling the same (or similar) functions. However, the findings reported here also suggest that these co-occurrences cannot simply be multimodal instantiations of several overlapping unimodal constructions. Given their high (statistical) co-occurrence, it seems unlikely that language users always construct these multimodal instantiations on the fly (as suggested by Hoffmann 2017). If that were the case, language use would be quite uneconomical. It seems more likely, at least in the case of *as if* constructions, that language users build crossmodal collostructions, i.e. strong links between different kinds of unimodal constructions (Uhrig 2018).

Apart from these substantiated conclusions, further, more tentative ones offer themselves. One of these is concerned with the different frequencies of occurrence when the features are considered. For instance, chunking the *as if* construction in one prosodic unit seems to be a central feature. Pitch, on the other hand, is a moderate predictor, while frowning is only a peripheral aspect of insubordinate *as if* constructions. The notion of crossmodal collostructions can explain these effects in terms of stronger and weaker associations between the individual constructions. And still, such a view can be complemented by Utterance Construction Grammar, which regards utterances as prototype categories with central and peripheral non-verbal associations. Given that the prototypical nature of verbal constructions has been argued for elsewhere, it seems reasonable to extend this conceptualization to the notion of multimodal constructions. The present article supports the idea that *as if* constructions have different multimodal profiles and that their features differ as regards cue validity. Therefore, seen as a multimodal gestalt with all features, each construction has a unique profile that enables the hearer to disambiguate the constructions in spoken English.

The present study is based on one family of constructions only and doesn't provide sufficient evidence for Utterance Construction Grammar. Still, the findings on *as if* constructions show that the grammar of spoken language should not be confined to the analysis of verbal elements alone, but that a multimodal perspective is worth consideration.

Author's address:

Department of Linguistics and Literature
English Linguistics
University of Bremen
Universitäts-Boulevard 13
28359 Bremen
Germany
claleh@uni-bremen.de

References

Attardo, Salvatore, Jodi Eisterhold, Jennifer Hay & Isabella Poggi. 2003. Multimodal markers of irony and sarcasm. *Humor – International Journal of Humor Research* 16(2), 243–60.

Barth-Weingarten, Dagmar. 2016. *Intonation units revisited: Cesuras in talk-in-interaction*. Amsterdam: John Benjamins.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.

Boersma, Paul & David Weenink. 2019. *Praat: Doing phonetics by computer*, version 6.0.52 [Windows]. Computer program.

Bolinger, Dwight. 1984. Intonational signals of subordination. *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, 402–13.

Bressem, Jana & Cornelia Müller. 2017. The 'Negative-Assessment-Construction' – A multimodal pattern based on a recurrent gesture? *Linguistics Vanguard* 3(s1). 20160053. https://doi.org/10.1515/lingvan-2016-0053

Brinton, Laurel J. 2014. The extremes of insubordination: Exclamatory *as if! Journal of English Linguistics* 42(2), 93–113.

Bryant, Gregory A. 2010. Prosodic contrasts in ironic speech. *Discourse Processes* 47(7), 545–66.

Bryant, Gregory A. & Jean E. Fox Tree. 2005. Is there an ironic tone of voice? *Language and Speech* 48(3), 257–77.

Caucci, Gina & Roger Kreuz. 2012. Social and paralinguistic cues to sarcasm. *Humor* 25(1), 1–22.

Cheang, Henry S. & Marc D. Pell. 2009. Acoustic markers of sarcasm in Cantonese and English. *The Journal of the Acoustical Society of America* 126(3), 1394–405.

Cheang, Henry S. & Marc D. Pell. 2011. Recognizing sarcasm without language: A cross-linguistic study of English and Cantonese. *Pragmatics & Cognition* 19(2), 203–23.

Cienki, Alan. 2017. Utterance Construction Grammar (UCxG) and the variable multimodality of constructions. *Linguistics Vanguard* 3(s1), 20160048. https://doi.org/10.1515/lingvan-2016-0048

Colston, Herbert L. 2020. Eye-rolling, irony and embodiment. In Angeliki Athanasiadou & Herbert L. Colston (eds.), *The diversity of irony*, 211–35. Berlin: De Gruyter Mouton.

Debras, Camille. 2017. The shrug: Forms and meanings of a compound enactment. *Gesture* 16(1), 1–34.

Debras, Camille. 2021. Multimodal profiles of *je* (*ne*) *sais pas* in spoken French. *Journal of Pragmatics* 182, 42–62.

Debras, Camille & Alan Cienki. 2012. Some uses of head tilts and shoulder shrugs during human interaction, and their relation to stancetaking. *Proceedings of the 2012 ASE/IEEE International Conference on Social Computing*, 932–7.

Du Bois, John W. 2007. The stance triangle. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 139–82. Amsterdam: John Benjamins.

Ekman, Paul & Wallace V. Friesen. 2003. *Unmasking the face*. Cambridge, MA: Malor Books.

Elff, Martin. 2022. *Package 'mclogit'. Multinomial Logit Models, with or without Random Effects or Overdispersion*, version 0.9.4.2. Computer program. http://mclogit.elff.eu/

Elvira-García, Wendy. 2019. Two constructions, one syntactic form: Perceptual prosodic differences between elliptical and independent clauses in Spanish. In Karin Beijering, Gunther Kaltenböck & María Sol Sansiñena (eds.), *Insubordination: Theoretical and empirical issues*, 240–64. Berlin: De Gruyter Mouton.

Elvira-García, Wendy, Paolo Roseano & Ana Ma Fernández-Planas. 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.

Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366–431. New York: Oxford University Press.

Feyaerts, Kurt, Geert Brône & Bert Oben. 2017. Multimodality in interaction. In
    Barbara Dancygier (ed.), *The Cambridge handbook of cognitive linguistics*, 135–56.
    Cambridge: Cambridge University Press.

Feyaerts, Kurt, Christian Rominger, Helmut Karl Lackner, Geert Brône, Annelies Jehoul,
    Bert Oben & Ilona Papousek. 2022. In your face? Exploring multimodal response patterns
    involving facial responses to verbal and gestural stance-taking expressions. *Journal of
    Pragmatics* 190, 6–17.

Ford, Cecilia E., Barbara A. Fox & Sandra A. Thompson. 2002. Constituency and the grammar of
    turn increments. In Cecilia E. Ford, Barbara A. Fox & Sandra A. Thompson (eds.), *The language
    of turn and sequence*, 14–38. Oxford: Oxford University Press.

Gibbs, Raymond W. 2000. Irony in talk among friends. *Metaphor and symbol* 15(1–2), 5–27.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalizations in language*.
    New York: Oxford University Press.

Goodwin, Charles. 2017. *Co-operative action*. Cambridge: Cambridge University Press.

Gries, Stefan Th. 2003. Towards a corpus-based identification of prototypical instances of
    constructions. *Annual Review of Cognitive Linguistics* 1(1), 1–27.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge
    University Press.

Hinnell, Jennifer. 2018. The multimodal marking of aspect: The case of five periphrastic auxiliary
    constructions in North American English. *Cognitive Linguistics* 29(4), 773–806.

Hoffmann, Thomas. 2017. Multimodal constructs–multimodal constructions? The role of
    constructions in the working memory. *Linguistics Vanguard* 3(s1), 20160042. https://doi.org/10.
    1515/lingvan-2016-0042

Huddleston, Rodney D. & Geoffrey K. Pullum *et al*. 2002. *The Cambridge grammar of the English
    language*. Cambridge: Cambridge University Press.

Imo, Wolfgang. 2007. Der Zwang zur Kategorienbildung: Probleme der Anwendung der
    Construction Grammar bei der Analyse gesprochener Sprache. *Gesprächsforschung–
    Online-Zeitschrift zur verbalen Interaktion* 8, 22–45.

Jehoul, Annelies, Geert Brône & Kurt Feyaerts. 2017. The shrug as marker of obviousness.
    Linguistics Vanguard 3(s1), 20160082. https://doi.org/10.1515/lingvan-2016-0082

Kaukomaa, Timo, Anssi Peräkylä & Johanna Ruusuvuori. 2014. Foreshadowing a problem:
    Turn-opening frowns in conversation. *Journal of Pragmatics* 71, 132–47.

Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Köhn, Arne, Timo Baumann & Oskar Dörfler. 2018. An empirical analysis of the correlation of
    syntax and prosody. *Proceedings of Interspeech 2018*, 2157–61. www.isca-speech.org/archive/
    pdfs/interspeech_2018/

Kreuz, Roger. 2020. *Irony and sarcasm*. Cambridge, MA: MIT Press.

Lehmann, Claudia & Alexander Bergs. 2021. As if irony was in stock: The case of constructional
    ironies. *Constructions and Frames* 13(2), 309–39.

Lelandais, Manon & Gaëlle Ferré. 2016. Prosodic boundaries in subordinate syntactic constructions.
    Paper presented at Speech Prosody 2016. https://hal.archives-ouvertes.fr/hal-01329167

Lelandais, Manon & Gaëlle Ferré. 2017. What do gestures in subordination tell us about clause
    (in)dependence? Paper presented at GESPIN2017-Gesture and Speech in Interaction.

Lelandais, Manon & Gaëlle Ferré. 2019. The verbal, vocal, and gestural expression of
    (in)dependency in two types of subordinate constructions. *Journal of Corpora and Discourse
    Studies* 2, 117–43.

López-Couso, María-José & Belén Méndez-Naya. 2012. On comparative complementizers in
    English: Evidence from historical corpora. In Nila Vázques (ed.), *Creation and use of historical
    English corpora in Spain*, 309–33. Newcastle upon Tyne: Cambridge Scholars.

Maschler, Yael. 2020. The insubordinate–subordinate continuum: Prosody, embodied action, and the emergence of Hebrew complex syntax. In Yael Maschler, Simona Pekarek Doehler, Jan Lindström & Leelo Keevallik (eds.), *Emergent syntax for conversation: Clausal patterns and the organization of action*, 87–125. Amsterdam: John Benjamins.

McNeill, David. 2005. *Gesture and thought*. Chicago: University of Chicago Press.

Ningelgen, Jana & Peter Auer. 2017. Is there a multimodal construction based on non-deictic *so* in German? *Linguistics Vanguard* 3(s1), 20160051. https://doi.org/10.1515/lingvan-2016-0051

Oxford English Dictionary. 2020. Oxford: Oxford University Press.

Perniss, Pamela. 2018. Why we should study multimodal language. *Frontiers in Psychology* 9, 1109.

Puts, David Andrew, Carolyn R. Hodges, Rodrigo A. Cárdenas & Steven J. C. Gaulin. 2007. Men's voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior* 28(5), 340–4.

R Core Team. 2019. *R: A language and environment for statistical computing*. Computer program. Vienna: R Foundation for Statistical Computing. www.R-project.org

Rockwell, Patricia. 2000. Lower, slower, louder: Vocal cues of sarcasm. *Journal of Psycholinguistic Research* 29(5), 483–95.

Rockwell, Patricia. 2007. Vocal features of conversational sarcasm: A comparison of methods. *Journal of Psycholinguistic Research* 36(5), 361–9.

Schoonjans, Steven. 2017. Multimodal Construction Grammar issues are Construction Grammar issues. *Linguistics Vanguard* 3(s1), 20160050. https://doi.org/10.1515/lingvan-2016-0050

Schoonjans, Steven. 2018. *Modalpartikeln als multimodale Konstruktionen: Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*. Berlin: De Gruyter.

Simpson, Paul. 2011. 'That's not ironic, that's just stupid': Towards an eclectic account of the discourse of irony. In Marta Dynel (ed.), *The pragmatics of humour across discourse domains*. Amsterdam: John Benjamins.

Steen, Francis & Mark B. Turner. 2013. Multimodal construction grammar. In Michael Borkent, Barbara Dancygier & Jennifer Hinnell (eds.), *Language and the creative mind*, 255–74. Stanford, CA: CSLI Publications.

Streeck, Jürgen. 2009. *Gesturecraft: The manu-facture of meaning*. Amsterdam: John Benjamins.

Szczepek Reed, Beatrice. 2011. *Analysing conversation: An introduction to prosody*. Basingstoke: Palgrave Macmillan.

Tabacaru, Sabina. 2019. *A multimodal study of sarcasm in interactional humor*. Berlin: De Gruyter Mouton.

Tabacaru, Sabina. 2020. Faces of sarcasm: Exploring raised eyebrows with sarcasm in French political debates. In Angeliki Athanasiadou & Herbert L. Colston (eds.), *The diversity of irony*, 256–77. Berlin: De Gruyter Mouton.

Tabacaru, Sabina & Maarten Lemmens. 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.

Uhrig, Peter. 2018. Multimodal constructions or crossmodal collostructions? Paper presented at the International Conference on Construction Grammar(s), Paris.

Uhrig, Peter. 2021. Multimodal corpus linguistics and machine learning methods for multimodal communication research. Paper presented at the 2021 International Conference on Multimodal Communication, Changsha.

Vigliocco, Gabriella, Pamela Perniss & David Vinson. 2014. Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences* 369(1651), 20130292.

Ward, Nigel G. 2019. *Prosodic patterns in English conversation*. Cambridge: Cambridge University Press.

Wehling, Elisabeth. 2017. Discourse management gestures. *Gesture* 16(2), 245–76.

Wells, J. C. 2006. *English intonation: An introduction*. New York: Cambridge University Press.

Wilson, Deirdre & Dan Sperber. 2012. Explaining irony. *Meaning and relevance*, 123–45. New York: Cambridge University Press.

Ziem, Alexander. 2017. Do we really need a Multimodal Construction Grammar? *Linguistics Vanguard* 3(s1), 20160095. https://doi.org/10.1515/lingvan-2016-0095

Zima, Elisabeth. 2017. On the multimodality of [all the way from X PREP Y]. *Linguistics Vanguard* 3(s1), 20160055. https://doi.org/10.1515/lingvan-2016-0055

Zima, Elisabeth & Alexander Bergs. 2017. Multimodality and construction grammar. *Linguistics Vanguard* 3(s1), 20161006. https://doi.org/10.1515/lingvan-2016-1006