

ARTICLE

# Multimodal-ish: prosodic and kinesic aspects of bounded and free uses of *ish*

Claudia Lehmann<sup>1</sup>  and Meike Pentrel<sup>2</sup>

<sup>1</sup>Department of English and American Studies, University of Potsdam, Potsdam, Germany; <sup>2</sup>Department of English and American Studies, University of Osnabrück, Osnabrück, Germany

**Corresponding author:** Claudia Lehmann; Email: [claudia.lehmann@uni-potsdam.de](mailto:claudia.lehmann@uni-potsdam.de)

(Received 07 October 2022; Revised 02 March 2023; Accepted 12 April 2023)

## Abstract

The article explores the prosodic and kinesic aspects of three different *ish* constructions using corpus data from the multimodal NewsScope Library of International Television News. The results reveal that bound *-ish* with ‘approximate’ meaning is longer in duration, higher in pitch, and shows more pitch variability than bound *-ish* with ‘properties’ meaning. Free *Ish* is also longer in duration and shows more pitch variability but is also prosodically set apart from its linguistic environment. Furthermore, the different *ish* constructions prove to be associated with different sets of kinesic features, although none of these reaches a significant level in the statistical model. It will be argued that the prosodic aspects mirror the constructional status of *ish*, whereas the kinesic aspects may be used to support their different functions.

**Keywords:** *ish*; multimodal construction grammar; prosody; gesture; quantitative corpus linguistics

## 1. Introduction

Communicative situations are predominantly multimodal (Bateman et al., 2017, pp. 7–9). When we talk to one another, we do not ‘just’ use words, but use intonation, a particular tone of voice, facial expressions, and manual gestures (to name but a few) for meaning-making. Therefore, spoken interactions are inherently multimodal (Feyaerts et al., 2017; Kendon, 2004; McNeill, 2005; Perniss, 2018; Vigliocco et al., 2014). Following this nowadays well-received insight, further questions offer themselves, which target the relationship between verbal, prosodic, and kinesic aspects of spoken communication: Are these aspects additive, that is, does one of the aspects support the other, or do they interact in such a way that mutual understanding is at stake if one is missing?

Multimodal construction grammar (MCxG) is a comparatively recent branch of Construction Grammar that investigates the relationship between verbal, prosodic, and gestural constructions and also explores the possibility of truly multimodal



constructions (Schoonjans, 2018; Zima & Bergs, 2017), that is, form-meaning pairings that include more than one mode (verbal, prosodic, kinesic, etc.) on the form side of the construction. A classic example for such a multimodal construction is the gestural use of deictic expressions such as *there* and *this/that* (Levinson, 1983, pp. 65–66). More recent research has shown that multimodal constructions extend beyond deictics. Two examples are the reactive *what-x* construction and its prosodic peculiarities (*What you find that appealing*; Pöldvere & Paradis, 2020) and ‘from beginning to end’, which is accompanied by a manual gesture in 8 out of 10 cases (Cánovas & Valenzuela, 2017), but there are quite a few others (e.g., Bressemer & Müller, 2017; Elvira-García, 2019; Gras & Elvira-García, 2021; Schoonjans, 2018; Zima, 2017). In addition to multimodal constructions, there is also good evidence for cross-modal associations between verbal and other kinds of constructions (see, e.g., Mittelberg, 2017; Uhrig, 2020). Even though they make weaker assumptions, they are of no lesser interest to MCxG.

In the present article, we report on an exploratory study concerned with the multimodality of the English morpheme *ish* in spoken language. Using a quantitative approach, we show that the different uses of *ish* introduced in Section 3 are accompanied by different prosodic and kinesic features. We also show that the prosodic features associated with the uses of *ish* mirror their constructional status and their level of entrenchment (lexicalized/non-lexicalized), whereas the kinesic features support the different meanings of *ish* in significant ways. Based on these empirical observations, we argue that the different *ish* constructions are, in fact, multimodal constructions.

## 2. Multimodal constructions and modes in spoken English

As mentioned above, we regard a construction to be multimodal if it is a form-meaning pairing that includes more than one mode. A question that naturally follows from such a definition is what a ‘mode’ is. Siefkes (2015) mentions two different uses of the term: on the one hand, ‘mode’ can be used in the sense of a sign system, that is, “a set of resources that often belong to a specific sign type and for which combination or application rules exist” (p. 114). On the other hand, it is used synonymously with sensory channel, “namely visual, auditory, haptic, olfactory, and gustatory” (p. 114). When assuming the second use of ‘mode’, a construction can only be multimodal when an auditory form (like [ðeə]) is combined with a visual form (like a pointing gesture) to signify a concept (like ‘a location at some distance from the speaker’), but it excludes combinations of morphosyntactic forms with suprasegmental forms (i.e., prosody) in spoken language since both make use of the same sensory channel. There is, however, some evidence that defining modes as sensory channels is problematic, to say the least. Bateman et al. (2017, pp. 114–115), for example, cite the McGurk effect (McGurk & MacDonald, 1976), which illustrates the effect the shape of the mouth has on our perception of sounds, to substantiate this claim. Given such an observable sensory overlap, the definition of ‘mode’ in terms of sensory channels does not seem useful.

If ‘mode’ is defined as a (potentially complex) sign system, a multimodal construction occurs when a morphosyntactic form is combined in spoken language with either a gesture and/or a prosodic form since there is abundant evidence that gestures (and other forms of proxemics) and prosody are independent sign systems. Bateman

(2016) argues that a semiotic mode is made up of three components, which are materiality, form, and discourse semantics. Gesture is a mode in spoken English given that it differs from spoken morphosyntactic constructions in all three respects. It uses a different materiality (hands rather than parts of the articulatory system), is different in form (e.g., pointing gesture vs. [ðeə]), and independently contributes to the discourse semantics (an expression like [ðeə] is hardly understood without the pointing gesture).

When mode is understood as a sign system, prosody is also an independent mode in spoken English. The materiality prosody uses partially overlaps with that of spoken morphosyntax. Both, for example, make use of the vocal folds, but while the tongue is important for articulating morphosyntactic units, it plays hardly any role for prosody. Vice versa, the diaphragm and the pressure it creates are highly important for prosody, but less so for the articulation process. Forms differ, too. While morphemes (made up of phonemes and stress placement) and their combinations (into words, phrases, idioms, etc.) are forms of interest in spoken morphosyntax, it is pitch, pitch movement, tempo, loudness (to name but a few), and their combinations that are of interest in prosody. Finally, the prosodic mode provides an independent contribution to the discourse semantics. Ward (2019), for example, introduced the notion of ‘prosodic constructions’, which “is a temporal configuration of prosodic features, has a meaning, is not necessarily closely aligned with words, can be present to a greater or lesser degree, can share aspects of meaning and form with related (sister and daughter) constructions, [and] can appear superimposed with other form-meaning mappings” (Ward, 2019, p. 108) and provides evidence for quite a few of these in English using a quantitative corpus approach. Empirical support for prosodic constructions also comes from Gras and Elvira-Garcia (2021), who find a distinct prosodic pattern for the in subordinate conditional construction in peninsular Spanish.

### 3. *Ish*<sup>1</sup>

The derivational suffix *-ish* has undergone a rather remarkable development recently. Originally a bound morpheme to create adjectives, it has gained autonomy and now exists also as a free morpheme. Free *Ish* can be used to qualify a previous utterance to indicate vagueness or to attenuate the proposition. According to the *Oxford English Dictionary Online* (OED, n.d.), it mainly functions “as a conversational rejoinder [meaning]: almost, in a way, partially, vaguely” (Oxford Languages, n.d.-b).

The suffix *-ish* attaches to various types of bases, including adjectives (*blueish* and *coldish*), adverbs (*soonish* and *early-ish*), (proper) nouns (*bookish* and *Obamaish*), compounds (*altar boy-ish*), phrases (*back-to-school-ish*), and numerals (*nine-ish*). Combinations with verbs seem to be limited to a couple of now lexicalized items such as *peckish* or *ticklish*. Being of Germanic origin, *-ish* (OE: *-isc*) was first used to derive adjectives describing ethnic belonging in Old English, as in *Engle-isc* or *frenc-isc* (Oxford Languages, n.d.-a). Such items can also be assumed to be lexicalized in present-day English. The combination with nominal bases can also be traced back to Old English, a prime example being *childisc* (see Oxford Languages, n.d.-a).

<sup>1</sup>To better distinguish between the different uses of *ish*, we will use the capital letter I for free *Ish*, whereas it will be noncapitalized when all kinds of uses are referred to.

From Middle English onward, we see an expansion concerning the potential bases for the suffix *-ish*, its meanings and functions. Meeting the prerequisites for the process of ‘debonding’ such as resemanticization, phonological strengthening, or flexibilization eases the way for the development of free *Ish* (Norde, 2009, p. 224, 2012). For more details on the historical development of bound *-ish*, see, for instance, Eitelmann et al. (2020) or Harris (2020, 2021).

In present-day English, the meanings of bound *-ish* are manifold but often related.<sup>2</sup> Most derivatives with *-ish* either express some association with the base, meaning that *X-ish* has (some) characteristics of *X*, as in *hippy-ish* and *librarian-ish*, or the addition of *-ish* indicates that *X* is almost like a set reference point, approximating that point on a scale, as in *wet-ish* or *old-ish* (see Traugott & Trousdale, 2013). Regarding the eligible adjectival bases, there seem to be intricate semantic regulations at work concerning, for instance, gradeability (see Bochnak & Csapak, 2014, for details).

Research on free *Ish* has mostly focused on its development, as an example of degrammaticalization (Norde, 2009) or constructionalization (Traugott & Trousdale, 2013; Trousdale, 2011), on its morphosyntactic features (Oltra-Massuet, 2017), as well as on semantic and occasionally pragmatic aspects (Bochnak & Csapak, 2014; Harris, 2021). So far, systematic corpus studies of *Ish* are the exception. This is due to its informal nature being mostly used in spoken language. Most discussions are based on the analysis of selected singular occurrences, taken from TV series, dictionary entries, or corpora. As an exception, Harris (2021) based her corpus study on Internet data, as *Ish* is used frequently on the Internet given the stylistic similarities of Internet data to spoken language. However, to the best of our knowledge, a systematic study of actual spoken data is still lacking. Prosodic features are sometimes mentioned, though. Bochnak and Csapak (2014, p. 440), for instance, observe that there seems to be a difference between bound *-ish* and free *Ish* stating that “Our evidence for this comes from our intuitions regarding the phonology of *-ish* versus ...*ish*: while propositional ... *ish* is always accompanied by a preceding pause, ordinary *-ish* and the use of precision-regulating *-ish* as applied to properties have the phonology of a bound morpheme.” Yet, again, no corpus study of actual spoken data confirms these intuitions.

Pragmatically, *Ish* can be used ‘as a sentence-final particle’ functioning as a degree modifier weakening a previous proposition (Bochnak & Csapak, 2014, p. 432). As free *Ish* is predominantly used in spoken conversation (Oxford Languages, n.d.-a), it is likewise found at the end of spoken utterances. This is also acknowledged by Bochnak and Csapak who state that *Ish* “states that a speaker is less than fully committed to an utterance” (2014, p. 448). Just like bound *-ish*, free *Ish* also indicates “a degree that is slightly less than the standard for the constituent it applies to” (Bochnak & Csapak, 2014, p. 433). An instance of free *Ish* in utterance final position modifying a proposition can be found in Examples (1) and (2).

<sup>2</sup>In the remainder of this article, we will exclusively use examples from our study (see below for details). Audio files for each numbered example can be accessed in the following repository (click the link or scan the QR code): <https://osf.io/ym6k7/>.



## (1) Happened-ish

01 TN well guess WHAT,  
 02 LAST night;  
 03 that' s eXACTly what happened;  
 04 ISH: .

## (2) Open-ish

01 MB well the happiest place on earth is (.) Open ish: .  
 02 <<creaky> aRIGHT.>  
 03 downtown disney PARTially reopened this morning;

It should be noted that *Ish* does not always modify the entire proposition, but it can also only modify the predicate (see Harris, 2021). *Ish* also “reduc[es] speaker commitment,” bearing similarities to discourse markers and hedges (Harris, 2021, p. 442). However, as Harris (2021) discusses, it cannot be entirely characterized as a discourse marker, as *Ish* “contributes meaning to the proposition, thereby altering it, while the common conception of discourse markers denies such propositional contribution” focusing solely on the pragmatic function (p. 442).

*Ish* can also appear in utterance-initial position, modifying the utterance of the previous speaker (see Trousdale, 2011) or as an answer to a question, as in Examples (3)–(5).

## (3) Right-ish

01 JC you WROTE the film;  
 02 you diRECTed the film;  
 03 and it took you (.) NINeteen years;  
 04 is that RIGH[ T to get-  
 05 EN [ ISH.  
 06 JC to[ get the movie made,  
 07 EN [yeah;  
 08 well I read the BOOK.

## (4) Play guitar-ish

01 ANON okay;  
 02 [ YOU can play guitar;  
 03 JF [ <<p> MANDolin;  
 04 yeah, <<creaky> I- I-> ISH;

## (5) Blind date-ish

01 RS you guys met on a BLIND date?  
 02 AW ((smacks lips))  
 03 GS we DID;  
 04 RS HOW[ was that;  
 05 AW [ISH;  
 06 I knew who HE was.

In Example (4), the first speaker notes that JF can play guitar (line 02). JF, however, does not commit fully to that statement; he agrees (‘yeah’), but he weakens speaker A’s assumption by employing *Ish* directly afterward (line 04). In Example (5), *Ish*

(line 05) serves as an answer to a closed interrogative (line 01), typically answered by *yes* or *no* or similar answers affirming or denying the proposition made in the question. GS affirms the proposition (line 03), whereas AW, answering after GS, modifies the idea that it was a blind date. She knew GS (line 06), so the set standard for a blind date was not completely met.

To sum up, we see a variety of uses of either bound *-ish* or free *Ish* in present-day English. Determining the base or rather the words in an utterance that are modified may be prone to misunderstanding. In her study of written corpus data, Harris (2021) points out that the way *-ish* is displayed orthographically, for example, in inverted commas, hyphenated, or set apart, can give “readers a clue as to what is modified” (p. 181). These types of orthographic variations seem to be especially frequent with the intermediate kind of *-ish/Ish*, as neither a traditionally bound morpheme nor as an instance of free *Ish* but rather located between these two extremes on a continuum (see Harris, 2021, p. 181). Often these kinds of intermediate *-ish/Ish* could theoretically still be attached to an immediately preceding base (as in Example (2)). However, the preceding pause and the fact that *ish* could often also be analyzed as modifying more than just the potential base (i.e., the preceding proposition) give these types of *ish* an oscillating status between free and bound (see also Pentrel, 2013).

To sum up, we see a variety of uses of either bound *-ish* or free *Ish* in present-day English, but, essentially, these can be boiled down to three schematic constructions:

1. Properties *-ish*: attaches to nouns and (noun) phrases and means ‘resembling/ having some properties of N’;
2. Approximate *-ish*: attaches to adjectives, adverbs, numerals, nouns, and phrases and means ‘approximating X’;
3. Free *Ish*: follows an utterance and modifies it in some way.

The review of previous research above shows that the morphological, syntactic, semantic, and pragmatic properties of *ish* have been looked at in some detail. To do so, the papers above mainly resort to (medially) written language as their empirical basis. The present article, in contrast, will analyze *ish* in (medially) spoken interactions (including both scripted and nonscripted examples) with a focus on its multimodal properties. More specifically, the article will show that the different uses of *ish* also differ regarding prosodic and gestural aspects of delivery, while the prosodic ones play a more prominent role. Eventually, the cognitive status of these features will be explained using an MCxG framework. To do so, we take a usage-based perspective on (spoken) constructions and also take the level of entrenchment of these constructions into account: As described above, *ish* is part of (at least) three schematic constructions, but also of dozens of other formally more fixed constructions, like *childish*, *clownish*, or *selfish*. These are so frequent that they need to be considered constructions in their own right (Goldberg, 2006). In the case of *selfish*, constructionhood is most obvious since its meaning is nonpredictable, being “concerned chiefly with one’s own advantage or welfare” (Oxford Languages, n.d.-c) rather than “resembling the self.” The way we dealt with this difficulty empirically will be explained in the following section.

## 4. Methods

### 4.1. The archive

To arrive at a sufficiently large empirical basis of medially spoken examples of *ish*, we used the *NewsScape Library of International Television News* (Steen & Turner, 2013). This archive is a large collection of televised discourse from various nations. In March 2021, it included almost 3 billion words of American English (Uhrig, 2021), and, given the fact that it is updated every day, is even larger by now. Since the archive contains recordings of discourses aired on television, the registers featured range on a continuum from scripted interactions (the news, TV series, speeches, and comedy routines) to semi-scripted (news interviews and late night show interviews) to only loosely scripted interactions (debates, discussions, and street interviews). The archive provides both audiovisual material and the corresponding captions, which enables searching for particular expressions.

### 4.2. Search procedure

We used the facilities provided by the Distributed Little Red Hen Lab to search for *ish* in the NewsScape Library, which resulted in a total of more than 5,000 hits. Given the fact that the present objective is to identify multimodal aspects of delivery, we removed any example in which the speaker's face was not visible. Technically, for annotating manual gestures, medium shots of the speakers are necessary (i.e., camera shots where the entire upper body of the speaker is visible), but since this was only the case for a small fraction of the hits, we decided to include hits with close-up shots (i.e., camera shots where only the face and, sometimes, the shoulders, but not the hands, are visible). This procedure allowed us to arrive at a substantial number of observations necessary for a solid quantitative analysis. We also removed duplicates and hits with considerable overlap between speakers. To arrive at a manageable size for manual annotations (see below), we further delimited our search to the past 3 years (Jul 2022 to Jul 2019). This procedure resulted in a total of 406 observations.

### 4.3. Annotation procedure

In the pilot phase of the study, the authors, independent of one another, looked at the first 50 observations in an informal way. After this step, they met and decided on an annotation scheme. As a result of this bottom-up approach, the level of detail of these annotations varies, depending on how promising the variables seemed after the pilot phase. The variables that were annotated using ELAN and Praat and their values as well as the abbreviations used for the statistical analyses in R are summarized in Table 1.

The only contextual variable that was annotated was the speaker of *ish* (SPEAKER). Since the NewsScape Library does not provide any speaker information in a systematic way, the speaker was identified manually. If they could not be identified, the speakers were labeled 'anonymous' and numbered consecutively. Textual variables that were annotated were the morphological status of *ish* (MORPHOLOGY), the level of entrenchment of the derivate (LEXICALIZATION), and the syntactic category of the base to which bound *ish* is attached (BASE). The morphological status of *ish* could take on the values 'free' or 'bound', depending on the syntactic category *ish* modified. When *ish* modified a syntactic category larger than a phrase, it was considered 'free';

Table 1. Annotation variables and values

<i>Contextual and textual variables</i>			
Variables	Abbreviations	Values	Abbreviations
SPEAKER	(none)	Adam Davidson, Adam Kinzinger, ..., Yasmin Vossoughian	(none)
MORPHOLOGY	(none)	bound, free	(none)
BASE	(none)	noun, adjective, adverb, numeral, phrase, other, none	N, Adj, Adv, Num, phrase, other, none
LEXICALIZATION	(none)	lexicalized, non-lexicalized	(none)
MEANING	(none)	property, approximation, modification	prop, approx, mod
ISH	(none)	property (lexicalized), property (non-lexicalized), approximation (lexicalized), approximation (non-lexicalized), free	prop(l), prop(nl), appr(l), appr(nl), free
<i>Prosodic variables</i>			
PAUSE_BEFORE_ISH	P1	(in ms)	na
PAUSE_AFTER_ISH	P2	(in ms)	na
DURATION	t	(in ms)	na
PITCH	F0	(in Hz)	na
STANDARD_DEVIATION	SD	(in Hz)	na
PROSODIC_INTEGRATION	(none)	integrated, distinct	(none)
MOVEMENT	(none)	level, falling, rising, rise-fall	(none)
<i>Kinesic variables</i>			
MANUAL_GESTURE		yes, no, NA	
HEAD		wiggle, nod, shake, tilt, turn, other	
WIGGLE_GESTURE		yes, no	
WIGGLE_ACTOR		hand, head, both, other, NA	
GAZE		somewhere, elsewhere, NA	
EYEBROWS		raised, frown, combination, other, none	
EYES		closed, upper lid raised, lower lid raised, cheeks raised, slit, other, none	
MOUTH		smile/laughter, nose wrinkle, lips pressed, other, none	

otherwise, it was categorized as ‘bound’. As the introductory section shows, the meaning of bound *ish* is largely affected by the syntactic category it is attached to, and so we also categorized the bases to which bound *ish* is attached, including the values ‘noun’, ‘adjective’, ‘adverb’, ‘phrase’, and ‘other’. Since free uses do not attach to a base, they were categorized as ‘none’. Using the base as the reference point, we distinguished between ‘property’, ‘approximation’, and ‘modification’ as values of MEANING. As laid out above, each *ish*-derivate is a possible construction and, therefore, we also annotated its presumed level of entrenchment. For that matter, we looked up the derivate in the OED (n.d.) and categorized it as either lexicalized (i.e., having an entry in






the OED) or non-lexicalized. Since free uses have an entry in the OED, all instances of free *Ish* were annotated as ‘lexicalized’. Due to this, we had to create a dummy variable we called ISH, whose values are all schematic constructions (non-lexicalized *ish* with ‘properties’ and ‘approximate’ meaning, respectively, free *Ish*) and the lexicalized constructions ending in bound *-ish*, further categorized according to their meaning. Treating all lexicalized *-ish* derivatives as separate categories (except for free *Ish*) would have resulted in an unmanageable size of constructions, and therefore we decided to assign them to one category.

To measure prosodic features, we used the speech analysis tool *Praat* (Boersma & Weenink, 2019) and manually measured possible pauses before (PAUSE\_BEFORE\_ISH) and after *ish* (PAUSE\_AFTER\_ISH), the duration of *ish* (DURATION), the mean pitch (PITCH), the standard deviation from the mean pitch (SD) as a measurement of possible pitch movements, the direction of this movement (MOVEMENT), and whether *ish* was integrated into the previous syntactic material or not (PROSODIC\_INTEGRATION). Prosodic integration was determined by a couple of factors. We considered *ish* to be prosodically integrated when it was not preceded by a pause and when it was either linked to a preceding consonant or vowel (i.e., when the syllable boundary was incongruent with the morpheme boundary, as, e.g., in *selfish*) or lacked a hiatus before a vowel (as in *Pollyannaish*). In all other cases, *ish* was considered to be prosodically distinct. Using *Praat* for acoustically analyzing data from non-laboratory environments is a delicate matter due to their noisiness. Therefore, we took great care in our measurements by setting the pitch ranges for each speaker individually and by excluding any datapoint that showed unusual pitch breaks.

The kinesic features were analyzed with the help of the annotation tool ELAN (ELAN, 2021). We used slow playback to annotate the presence of a manual gesture (MANUAL\_GESTURE), embodied actions we labeled ‘wiggles’, that is, quick back-and-forth movements on an axis (WIGGLE\_GESTURE), and the part of the body performing the wiggle (WIGGLE\_ACTOR). We found the hands (one or both) and the head to be the most important actors, labeling all other actors ‘other’. If the hands were not sufficiently visible (as in close-up shots), this was labeled NA. For the first 200 observations, manual gestures other than wiggles were also annotated. However, these proved to be quite heterogenous, which is why they have not been annotated for the entire dataset and will not be reported here. We also annotated head movements (HEAD) and the gaze direction of the speaker on *ish* (GAZE). Since gaze direction is an imprecise measurement when done manually, we largely distinguished between looks directed at ‘somewhere’ (i.e., looks to the camera, the recipient, or the audience) and ‘elsewhere’ (i.e., looks that avoided eye contact with a (virtual) person, including the camera). When the camera shot during the utterance of *ish* was a close-up, we considered the sequential context to determine the gaze direction. If this was not possible, we labeled this NA. Moreover, we annotated movements in the eyebrow (EYEBROW), eye (EYE), and mouth regions (MOUTH). To annotate facial actions, we used a subset of the Facial Action Coding Manual proposed by Ekman and Friesen (1978). We also used the zoom function in ELAN to get a detailed shot of the speaker’s face. To illustrate this, a sample of each facial variable (made by the same speaker) can be found in Table 2. If the facial expression could not be annotated after zooming in, we labeled this NA.

**Table 2.** Sample annotations for facial variables

Variables	Values	Screenshots
EYEBROW	'raised'	
EYE	'slit'	
MOUTH	'smile'	

**Table 3.** Intercooder reliability for the double-coded variables

Variables	Estimates (Cohen's $\kappa$ )	Lower	Upper
BASE	0.76	0.69	0.83
MORPHOLOGY	0.67	0.52	0.83
WIGGLE_GESTURE	0.60	0.49	0.72
WIGGLE_ACTOR	0.61	0.42	0.79
HEAD	0.37	0.29	0.46

To ensure reliability, the two authors of this article annotated some of the variables for a subset of 200 instances independent of one another. These variables were MORPHOLOGY, BASE, WIGGLE\_GESTURE, WIGGLE\_ACTOR, and HEAD. The *psych* package (Ravelle, 2022) in the statistical program R (R Core Team, 2019) was used to calculate Cohen's kappa for estimating the intercoder reliability. For many of the variables, the intercoder reliability was moderate, only for head movements agreement was minimal. The details can be found in Table 3.

The authors discussed all disputed cases until agreement was reached. After this step, the first author annotated the remaining datapoints according to the annotation scheme that was agreed upon.

## 5. Statistical analysis

The statistical analysis was done in R (R Core Team, 2019). We fitted a generalized linear mixed-effects model because of its flexibility and statistical power. It allowed us to assess the relationship between the use of *ish* and both discrete and continuous variables (and their combinations). Furthermore, it allowed us to include random variables (i.e., variables that are allowed to vary independently). More specifically, we fitted a polytomous model with MEANING as the dependent variable (with the 'properties' function<sup>3</sup> as the reference level) using the function *mblogit* of the *mclogit* package (Elff, 2022). To do so, numeric variables were centered, except for PITCH, which was log-transformed. The initial model was an intercept-only model with LEXICALISATION and SPEAKER as random effect terms. Due to the high number of different speakers in the dataset ( $N = 257$ ), problems with convergence occurred and, given that, we decided to exclude SPEAKER as a model term. This is, of course, not unproblematic because the dataset is not balanced for SPEAKER and some of the speakers, who occur more frequently than others, might skew the results. We will discuss this point in Section 5. Next, we added prosodic and kinesic variables as fixed effects to the model one at a time. We used the Akaike information criterion (AIC) to assess the model fit, and we kept any term in the model that resulted in a lower AIC. When the AIC was only marginally higher than in the previous model (i.e., the difference was lower or equaled 5), we will mention this below. The summary of the final model was made with the *tab\_model* function of the *sjPlot* package (Lüdtke, 2021). The plots showing the interaction and mosaic plots between the variables of interest were made using the *interaction.plot* and *mosaicplot* functions in base R. The dataset, the R script, and all figures can be accessed here (click the link or scan the QR code): <https://osf.io/ym6k7/>.



<sup>3</sup>Note that we use *meaning* and *function* synonymously here.

**Table 4.** Summary of the fitted model for the meanings of *ish* and their multimodal features they are accompanied with

MODEL INFO			
Observations: 304			
$N_{\text{lexicalisation}}$ : 2			
Dependent variable: MEANING			
Type: mixed effects generalized linear regression			
Error distribution: multinomial			
Link function: logit			
MODEL FIT			
AIC = 432.14, BIC = 506.48			
approximation vs. properties			
Predictor	Estimates	CI	<i>p</i>
(Intercept)	-5.13	-10.47 to 0.22	0.060
Duration	0.62	0.12 to 1.12	<b>0.015</b>
Pitch	1.24	0.22 to 2.27	<b>0.018</b>
Standard deviation	-0.50	-0.89 to -0.11	<b>0.012</b>
Prosodic integration	-0.42	-1.15 to 0.30	0.253
Pause before <i>ish</i>	0.56	-0.11 to 1.24	0.102
Pause after <i>ish</i>	-0.31	-0.69 to 0.08	0.118
Rising pitch movement	-0.96	-1.97 to 0.05	0.063
Directed gaze	-0.18	-0.86 to 0.50	0.599
Slit eyes	0.28	-0.65 to 1.22	0.550
modification vs. properties			
Predictor	Estimates	CI	<i>p</i>
(Intercept)	-10.45	-22.67 to 1.77	0.094
Duration	1.16	0.33 to 1.99	<b>0.006</b>
Pitch	1.69	-0.50 to 3.88	0.129
Standard deviation	-0.81	-1.45 to -0.17	<b>0.013</b>
Prosodic integration	-2.95	-4.79 to -1.11	<b>0.002</b>
Pause before <i>ish</i>	0.61	-0.18 to 1.41	0.130
Pause after <i>ish</i>	-0.10	-0.81 to 0.61	0.789
Rising pitch movement	-0.24	-2.24 to 1.75	0.811
Directed gaze	-0.12	-1.58 to 1.34	0.874
Slit eyes	1.51	-2.31 to 5.33	0.438

Note: Significant variables are highlighted in bold face.

## 6. Results

The fitted model is summarized in Table 4.

Table 4 shows that the ‘approximation’ function of *ish*, when being compared with the ‘properties’ function, shows significant differences regarding its duration, its mean pitch height, and its standard deviation. Moreover, rising pitch movement reached borderline significance. Prosodic integration, the use of pauses before and after the *ish*, gaze direction, and slit eyes improved the model fit, but did not reach a significant level. The ‘modification’ meaning of *ish* can significantly be distinguished from the ‘properties’ function by its duration, its standard deviation from the mean pitch, and its prosodic (dis)integration. Mean pitch height, the uses of pauses before and after *ish*, pitch movement, gaze direction, and slit eyes, while improving the

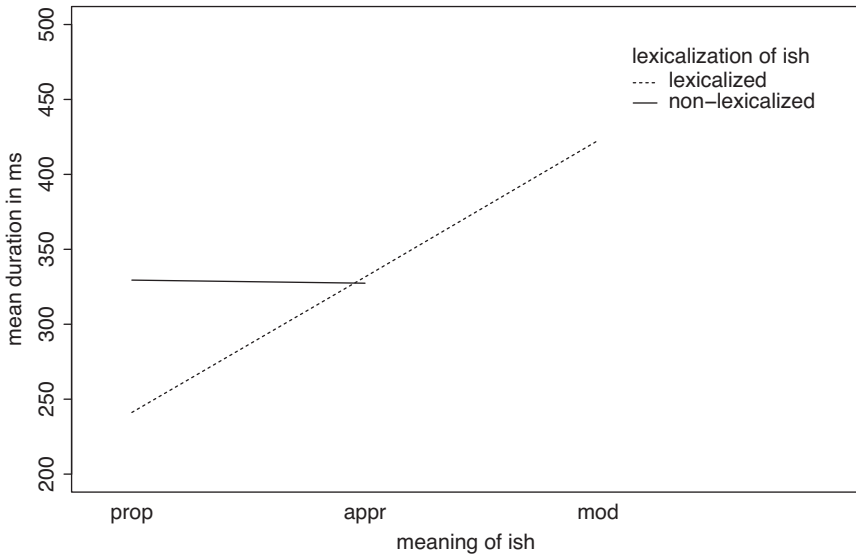


Figure 1. Mean duration (in milliseconds) for the lexicalized and non-lexicalized uses of *ish*.

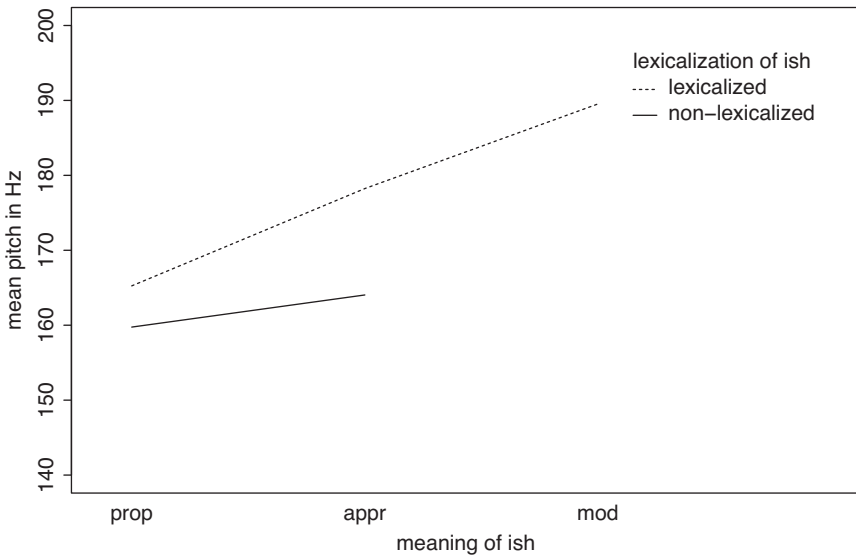
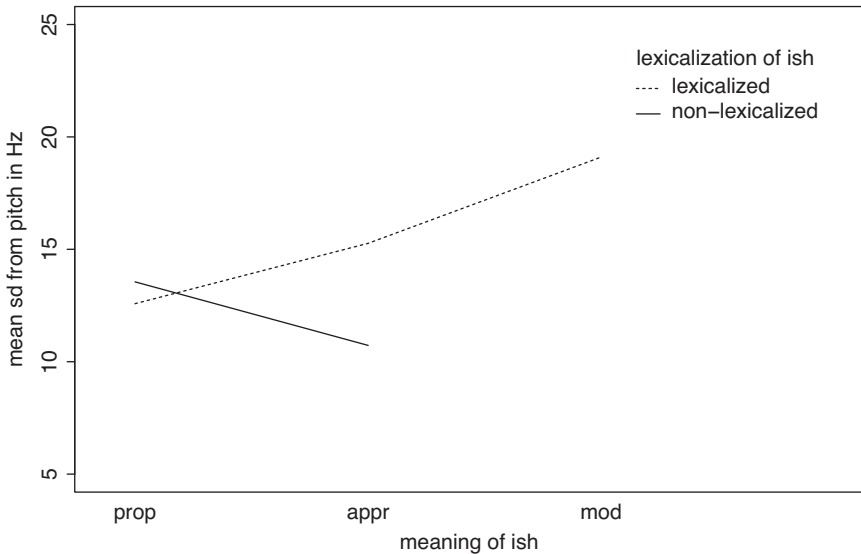


Figure 2. Mean pitch (in hertz) for the lexicalized and non-lexicalized uses of *ish*.

model, did not reach a significant level to sufficiently distinguish the modification function from the properties function.

Figs. 1–3 provide interaction plots for duration, mean pitch, and the standard deviation from the mean pitch for the lexicalized and non-lexicalized uses of *ish*.



**Figure 3.** Mean standard deviations from the mean pitch (in hertz) for the lexicalized and non-lexicalized uses of *ish*.

Figs. 1–3 illustrate that non-lexicalized uses of *ish* with either meaning have about the same duration. Lexicalized uses, on the other hand, differ in duration, depending on their meaning: when *ish* is attached to a noun and has the meaning ‘having the properties of N’ (e.g., in *clownish*), it is rather short, whereas it is longer when being attached to other kinds of bases, having an approximating function (as in, e.g., *soonish*). It is longest in duration when used with modifying function (i.e., free *Ish*). Regarding mean pitch, lexicalized uses of *ish* are higher in pitch than non-lexicalized uses. In addition, the pitch also differs for the three meanings. The ‘modification’ function is higher in pitch than the ‘approximating’ function, which, in turn, is higher in pitch than the ‘properties’ function. These results need to be treated with great caution, though. Mean pitch is a variable that is highly speaker-dependent, given the fact that voices vary with the size of the larynx (first described as the Frequency Code in Ohala, 1983). Since we could not include the speaker in the model, this result might be skewed. Concerning the standard deviation from the mean pitch, the interaction plot suggests that non-lexicalized *ish* with ‘approximate’ meaning (as, e.g., in *normalish*) shows the smallest variation in pitch movement, followed by the ‘properties’ meaning, irrespective of the level of lexicalization. Lexicalized *ish* with ‘approximate’ function (*soonish*) shows more variation in pitch than its non-lexicalized counterpart (*normalish*), and free *Ish*, having a modifying function, shows the greatest pitch variation.

Figs. 4 and 5 illustrate the associations between the different uses of *ish* and their prosodic integration into the previous linguistic material (Fig. 4) and pitch movements (Fig. 5).

Fig. 4 shows that *ish* having the ‘properties’ function tends to be prosodically integrated to its base when it is lexicalized (*clownish*), whereas it can be either

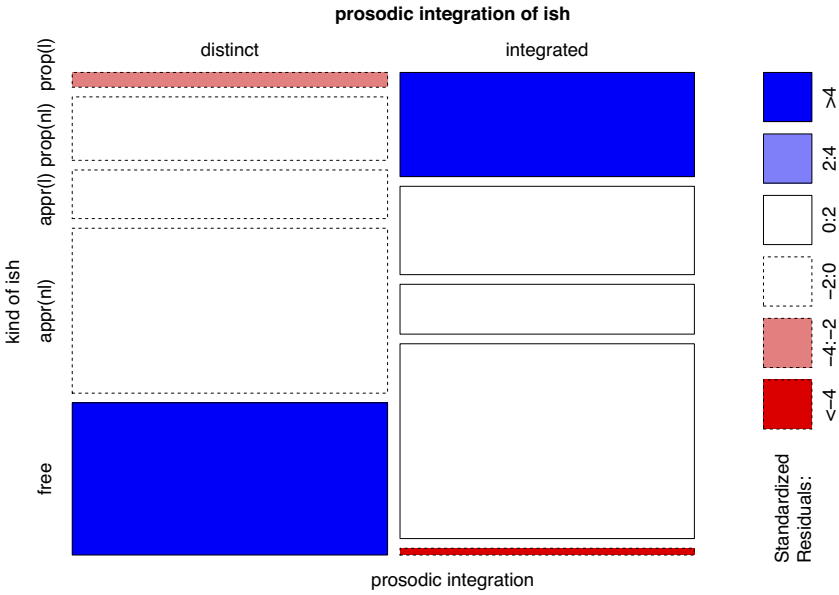


Figure 4. Mosaic plot illustrating the relative frequencies between the uses of *ish* and their prosodic integration.

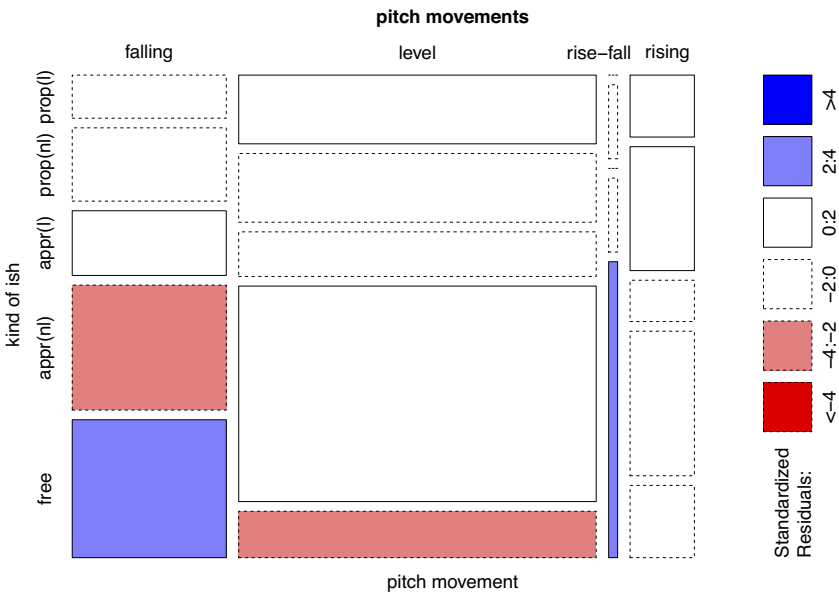


Figure 5. Mosaic plot illustrating the relative frequencies of the different uses of *ish* and their pitch movements.

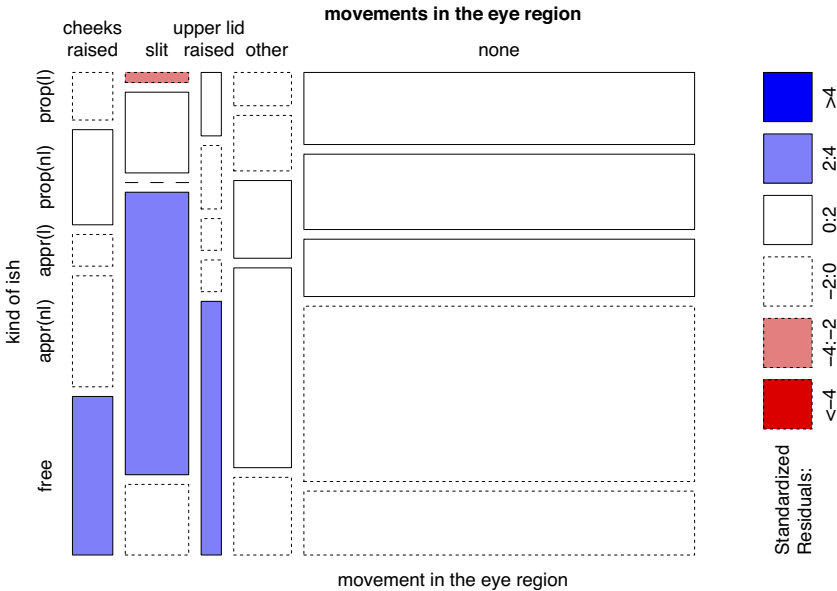


Figure 6. Mosaic plot illustrating positive (blue) and negative (red) associations between the different uses of *ish* and movements in the eye region.

integrated or distinct from its base when it is non-lexicalized (as in, e.g., *old-guy-ish*). *Ish* having an ‘approximate’ function shows no tendency, be it lexicalized or non-lexicalized when compared to the other uses of *ish*. Free *Ish* shows a strong tendency for being prosodically distinct. As regards pitch movement (Fig. 5), the plot shows that *ish* with ‘properties’ function has no preference for any kind of pitch movement, neither its lexicalized nor its non-lexicalized variant. *Ish* with ‘approximate’ meaning, on the other hand, shows a slightly negative association with falling pitch when it is non-lexicalized (*normalish*), but no preference for any other kind of movement. Free *Ish* with modifying function has a general tendency for falling pitch movements, including rise-falls.

Fig. 6 illustrates the associations between the different uses of *ish* and movements in the eye region.

Fig. 6 shows that lexicalized *ish* with ‘properties’ meaning (*clownish*) shows a slightly negative tendency for slit eyes, whereas its non-lexicalized counterpart (*old-guy-ish*) shows no tendency at all. *Ish* with ‘approximate’, non-lexicalized meaning (*normalish*), on the other hand, shows a slight tendency for the eyes to be slit, whereas no tendency for its lexicalized counterpart (*soonish*) can be observed. Free *Ish* with modifying function is accompanied by either raised cheeks or upper eyelids.

The gaze direction observed for the different kinds of *ish* shows some interesting tendencies, but none of these reached a significant level. Therefore, gaze direction will not be considered any further here.

In addition to the terms that entered the final model, there were quite a few variables whose inclusion in the model did neither improve nor worsen the model fit.



**Table 5.** Summary of the observations made for the different uses of *ish* and some selected kinesic features

	'properties'		'approximate'		'modification'
	Lex.	Non-lex.	Lex.	Non-lex.	
EYEBROWS	(none)	(none)	(none)	(none)	(none)
MOUTH	(none)	(none)	(none)	(none)	positive for 'smile'
HEAD	(none)	(none)	(none)	positive for 'wiggle', negative for 'tilt'	positive for 'tilt', negative for 'wiggle'
MANUAL_GESTURE	negative for 'yes'	(none)	(none)	(none)	(none)
WIGGLE_GESTURE	negative for 'yes'	(none)	(none)	positive for 'yes'	(none)

These were movements in the eyebrow and mouth regions, head movements, the presence of a wiggle, and the presence of a manual gesture. For reasons of space, the mosaic plots for these variables will not be shown here, but they are available in the repository linked above. Still, there are some interesting observations to be made, which are summarized in Table 5.

## 7. Discussion

The model reported above suggests that the prosodic features alone are significant in distinguishing the meanings of *ish* and, as a consequence, in distinguishing the different *ish*-constructions. Therefore, we will first discuss these using examples to illustrate the uses of *ish* and their prosodic features and then turn to discuss the kinesic features.

### 7.1. Prosodic features

To discuss the prosodic features, we selected examples uttered by three male speakers (abbreviated CC, JF, and JT, respectively). This allows us to discuss the prosodic features of the constructions with only a minimal influence of the speaker (gender) as a confounding variable on, for instance, the mean pitch.

Examples (6) and (7) are both illustrating the 'properties' meaning of *ish*. Example (6) is an example of a lexicalized variant of the 'properties' function; the word *Pollyannaish*, meaning "resembling Pollyanna; naively cheerful and optimistic; unrealistically happy" (Oxford Languages, n.d.-d). Example (7) is an example of the non-lexicalized variant; *slawish* is an ad hoc creation used in the sense of 'resembling slaw'.

(6) Pollyannaish (CC)

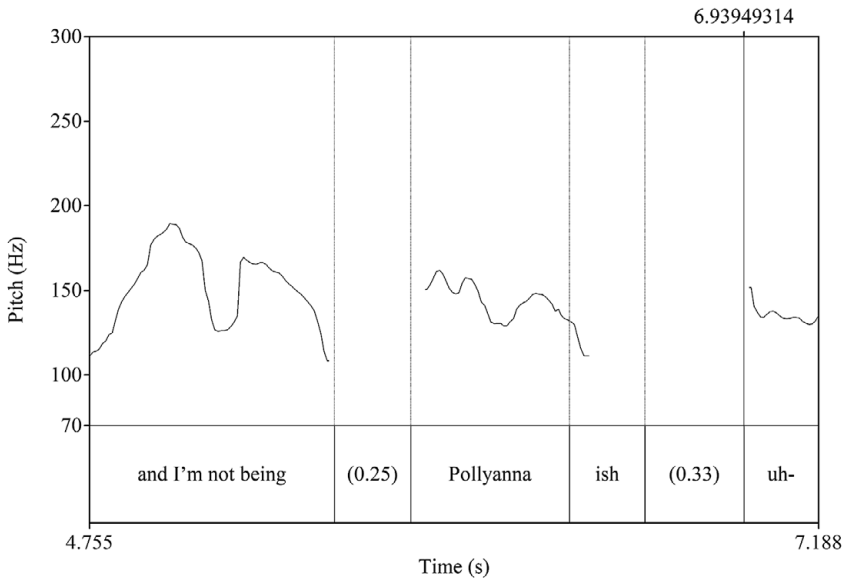


Figure 7. Pitch movements of Example (6).

(7) Slawish (JF)

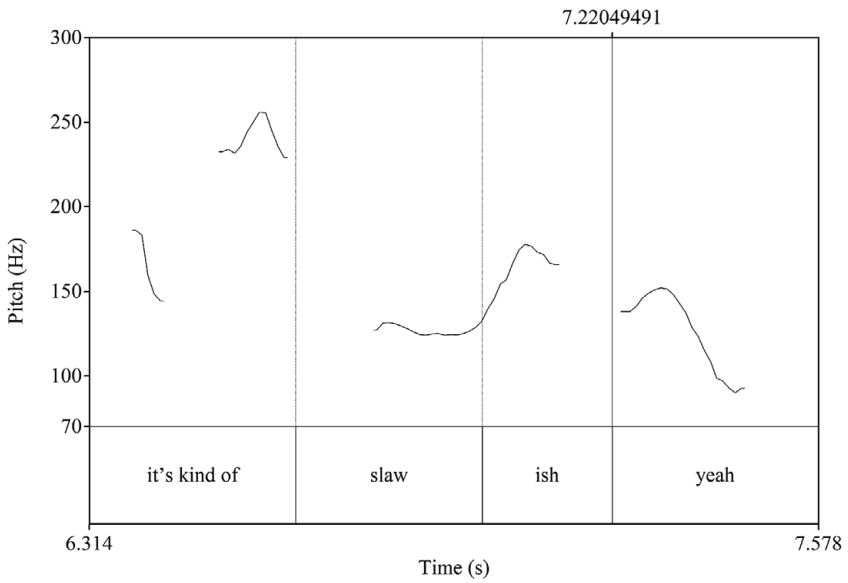


Figure 8. Pitch movements of Example (7).

Figs. 7 and 8 illustrate the prosodic features of *ish* having the ‘properties’ function. They show that ‘properties’ *ish* is comparatively short (with the *ish* in (6) being 280 ms and in (7) 212 ms long). Neither of the two uses is preceded by a pause, and both are prosodically integrated (indicated by the continuous lines). In addition, both examples are within the normal pitch range of the respective speaker. Example (6) has a mean pitch of about 128 Hz, which might also be due to the fact that it occurs at the end of the prosodic unit and the pitch tends to drop in this environment (at least in standard American English; see, e.g., Barth-Weingarten, 2016). Since the falling of the pitch already started before the onset of *ish*, the standard deviation from the mean pitch in Example (6) is rather small nonetheless, with 9 Hz. The linguistic environment being the end of a prosodic unit might also explain why Example (6) is followed by a (short) pause, even though the model reported in Section 4 does not suggest this. Example (7) is higher in pitch than (6), with 165 Hz, but compared to the surrounding pitch, this seems to be within the normal range of the speaker, given the fact that the rise on ‘kind of’, shortly before, is much higher. Moreover, with a standard deviation from the mean pitch of 11 Hz, the rise in pitch is not overtly great here.

In contrast to these, examples (8) and (9) illustrate lexicalized and non-lexicalized constructs of ‘approximate’ *ish*, respectively. The model reported above predicts that ‘approximate’ *ish* is longer in duration and higher in pitch. In addition, lexicalized approximate *ish* is predicted to show more pitch variability, whereas non-lexicalized approximate *ish* shows the opposite tendency.

(8) trueish (CC)

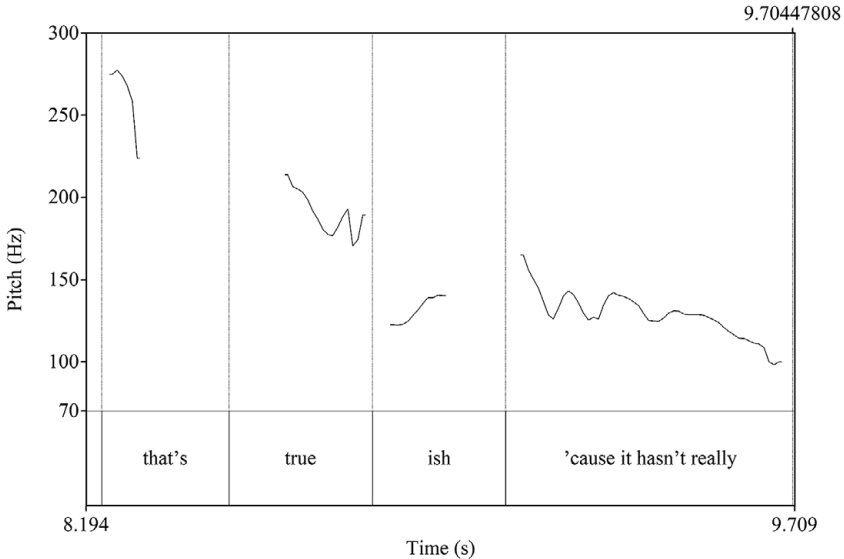


Figure 9. Pitch movements for Example (8).

## (9) normalish (JT)

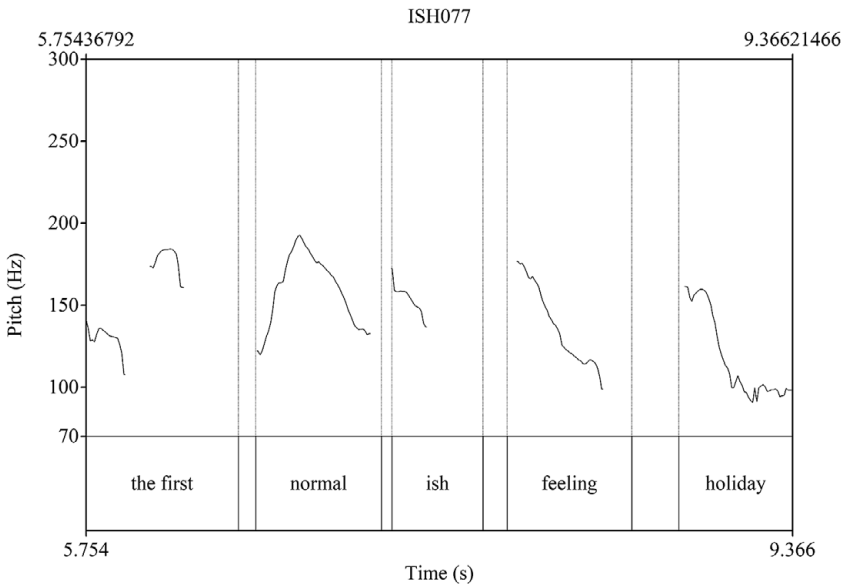


Figure 10. Pitch movements of Example (9).

Examples (8) and (9) illustrate that ‘approximate’ *ish* is, on average, longer in duration than ‘properties’ *ish*, with a duration of 247 ms and 449 ms, respectively. Even though Example (8) is comparable in length with Examples (6) and (7), Example (9) is much longer in duration and, thus, exemplifies the range the duration of ‘approximate’ *ish* can have. Examples (8) and (9) are significantly higher in pitch than Examples (6) and (7), with a pitch of 135 Hz and 154 Hz, respectively. Even though Example (8) seems low in pitch at first sight, it is higher than Example (6), which was produced by the same speaker, CC, and, thus, serves as a good reference point. In Example (9), the pitch used for *ish* is also comparatively high when compared with the surrounding pitches, which are about the same level. Regarding pitch variability, Examples (8) and (9) illustrate the opposite tendencies for lexicalized and non-lexicalized variants of ‘approximate’ *ish*. (8) is an example of lexicalized *ish* and has a standard variation of about 17 Hz from its mean. This movement is clearly audible (and visible). In contrast, Example (9) shows less pitch movement with a standard deviation of about 10 Hz, which is comparable to the standard deviations observed for ‘properties’ *ish*. In addition to these observations predicted by the model, Examples (8) and (9), in contrast to Examples (6) and (7) also illustrate that ‘approximate’ *ish* can (but need not be) prosodically distinct from its base. The distinctness here manifests itself by a sudden pitch upstep in Example (8) and a downstep in Example (9) (Figs. 9–10).

Free *Ish* is illustrated in Example (10).

(10) fair game ish

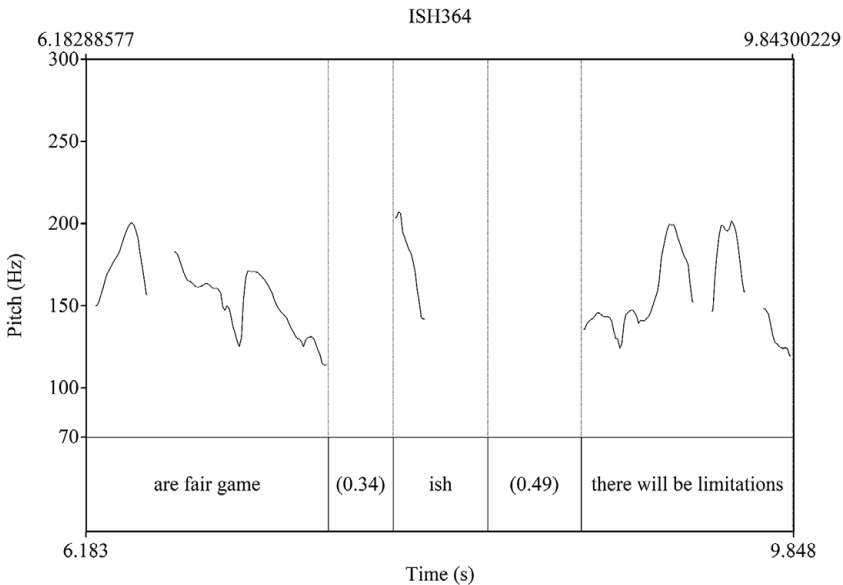


Figure 11. Pitch movements of Example (10).

The model predicts that free *Ish* is longer in duration, shows more pitch variation, and is prosodically distinct when compared with ‘properties’ *ish*. All these features are present in Example (10). With a duration of 426 ms, it is rather long, and with a standard deviation of 23 Hz, the (falling) pitch movement is clearly visible (and audible). In contrast to both ‘properties’ and ‘approximate’ *ish*, free *Ish* is always distinct from the linguistic material it modifies. In Example (10), this distinctness is achieved with the help of pauses preceding and following the *ish* plus a pitch upstep (of about 90 Hz).

Examples (6)–(10) have illustrated the prosodic properties of the different uses of the morpheme *ish*. Given their significance in the model, there is substantial evidence that these features are integral parts of the constructions that not only formally distinguish them, but also support their individual meanings. ‘Properties’ *ish*, being treated as the reference construction here, suggests a comparatively high commitment to the epistemic stance of the speaker to what is uttered: In Example (6), the speaker commits to the claim that he is not resembling Pollyanna, and in Example (7), the speaker commits to the claim that some entity (here: his hair) resembles slaw regarding its consistency. This meaning, and in particular, its commitment to the truth value of this meaning, manifests itself in being uttered in a short, unmarked, and integrated way. Since there is, normally, no need to do so, ‘properties’ *ish* is usually not set apart from the surrounding linguistic material.

This is different for ‘approximate’ *ish*. ‘Approximate’ *ish* does not fully commit to the truth value of the utterance it is part of but indicates a tentative commitment. This

tentativeness is supported by its prosodic features. It is longer in duration because the speaker either needs more time to think about an appropriate expression or constructs this to be the case. Since it needs to be made prominent to some extent, it is higher in pitch and might also be prosodically distinct. The level of lexicalization of ‘approximate’ *ish* seems to play a role, though, at least regarding pitch movement. Non-lexicalized constructs of ‘approximate’ *ish* show less pitch variation and a tendency for non-falling pitch movements. Falling pitch movements (on entire intonation units) often signal definiteness (Wells, 2006), and thus it seems plausible to assume that avoiding a non-falling pitch supports the lack of epistemic certainty. However, it could be argued that, since *ish* is just a morpheme, and ‘approximate’ *ish* cannot constitute an intonation unit by itself (but is only part of one), pitch movement is not an applicable category. While it is certainly true that bound *ish* needs to be part of some other linguistic material, the results of the different standard deviations and directions of the pitch movements need to be accounted for. We suggest that both lexicalized and non-lexicalized variants of ‘properties’ *ish* and the lexicalized variant of ‘approximate’ *ish* show no tendency for a particular pitch movement exactly because they are part of a larger intonation unit and are integrated into the larger pitch movements of this unit. Non-lexicalized ‘approximate’ *ish*, on the other hand, tends to be uttered with level pitch because the speaker indicates their tentative commitment to the resulting construct, which needs to be constructed by the speaker and deconstructed by the recipient(s), because the construct is not readily available to them as a construction. Considering the usage-based commitment we set out in the beginning, lexicalized variants of ‘approximate’ *ish* are likely better treated as independent constructions, whereas the non-lexicalized variants are genuine examples of the constructions.

As argued in Section 2, free *Ish* is used to qualify an immediately preceding utterance. As such, it constitutes an utterance itself in most of the cases and this utterance is realized as one intonation unit. This intonation unit is longer in duration and, often, preceded and followed by pauses, when compared with the bound uses of *ish*, presumably because the unit it attaches to and, consequently, qualifies, is larger. In using more time, two effects are achieved: on the one hand, the hearer is granted more time to arrive at possible implications of the modification, and, on the other hand, the speaker puts extra emphasis on the modification.

## 7.2. Kinesic features

The results reported in Section 5 have shown that none of the kinesic features reached a significant level in the model, which suggests that these are not integral parts of the constructions. However, mosaic plots (based on chi-square statistics) suggest some interesting associations between the constructions and movements in the eye and mouth regions, head movements, and the use of (manual) gestures. These will be illustrated in what follows. Since both non-lexicalized uses of ‘properties’ *ish* and lexicalized uses of ‘approximate’ *ish* showed no associations with kinesic features whatsoever, these will not be illustrated here. We used the conventions proposed in Mondada (2018) to transcribe the examples multimodally.

The first example, Example (11), illustrates the features of lexicalized ‘properties’ *ish*.

(11) Childish

1	speech	Second;
	head	+tilt to the right ----->
2	speech	to #stop being so CHIL#dish (.) uh with china;
	head	--+nod+-----+nod+-----+nod+-->
	fig	#Fig. 12                      #Fig. 13



Figure 12.



Figure 13.

This example illustrates the kinesic features accompanying the use of lexicalized variants of ‘properties’ *ish*. In fact, some kinesic properties can be observed in this example, but none of these seem to be triggered by or associated with the meaning of *-ish*. When continuing his line of argumentation beginning with ‘second’, the speaker tilts his head to the right, an action that has been reported to be used when speakers disaffiliate with a third party (Debras, 2017; Debras & Cienki, 2012). This seems to be the case here since the speaker criticizes some other person(s) for being ‘childish’ and the head remains in a slightly tilted position when the speaker continues (see Figs. 12 and 13). Furthermore, he slightly nods each time he utters a syntactic head in the prosodic unit ‘childish’ is part of (‘stop’, ‘childish’, and ‘China’), including a nod on ‘childish’ itself. Thus, the nods fulfill the function of beat gestures (see McNeill, 1992, p. 15) and do not seem to fulfill any other function. Apart from these head movements, the speaker does not use any further kinesic features.

The non-lexicalized use of approximate *ish*, in comparison, is, often, accompanied by slit eyes, a wiggle gesture (performed by either the head or the hands), but no head tilts. This is illustrated in Example (12).

## (12) Twenty percent ish

- 1 speech SO: uhm- #  
 eyes \*blinks frequently----->  
 hands Δboth hands folded and raised to chestΔ  
 fig #Fig. 14
- 2 speech ROUGHly #twenty percent ish #h° of the country;  
 eyes -----\* \*slit-----\*  
 head +down-----+ +wiggle-----+  
 hands Δboth hands wiggle-----Δ  
 fig #Fig. 15 #Fig. 16
- 3 speech has at least ONE dose of the vaccine;  
 eyes \* lower lid raised----->>



Figure 14.



Figure 15.



Figure 16.

In Example (12), the speaker tries to give an estimate of the number of U.S. citizens who received their first shot of a coronavirus vaccine and hedges this estimate, first, by prefacing it with the adverb ‘roughly’ and with the *ish* following it. Since *ish* in this example is phrase internal (both regarding the syntactic and intonation phrase), we classified it as a bound *ish* with ‘approximate’ meaning. What is striking here is the speaker’s use of wiggle gestures, using both the hands and her head. Before she gives the estimate, she hesitates and plans the following utterance(s). This becomes obvious not only by the use of the filled pause (‘uh’), but also by the frequent blinks, which indicates mental load (Holland & Tarlow, 1972). In addition, she raises both hands to chest height and folds them (see Fig. 14). In this context, it seems as if this gesture is also used to prepare the gesture that follows: when finally delivering the estimate, the speaker unfolds her hands and starts wiggling them on a sagittal axis with her fingers spread. This manual wiggling continues until the end of the syntactic and intonational phrase. In addition to this, she also lowers her head and starts wiggling it on the onset of *ish* and stops doing so after the outbreak that follows. Thus, while the entire phrase is accompanied by a manual wiggle, *ish* is further accompanied by a head wiggle, lending strong support to the assumption that it is the approximate *ish* construction that is associated with a wiggle gesture. Moreover, the speaker narrows her eyes to a slit (see Figs. 15 and 16) during the utterance.

Both the wiggle gestures and the eye slit support the meaning of the approximate *ish* construction. The wiggle gesture, be it by the hand(s) or the head, is performed by quick and small movements around an axis. When used with approximate *ish*,



the speaker tries to come close to some entity (e.g., a property or a number) and this metaphorical action of nearing an appropriate term is embodied with the help of an oscillating movement of some body parts, mainly the hands or the head. The midpoint of the simulated axis on which the wiggle is performed is the imaginative location of that entity. Similarly, narrowing one's eyes can serve the purpose to see an entity more clearly. In the case of approximate *ish*, the entity is not physically present, but still the speaker seems to simulate this experience to get closer to this entity.

Finally, Example (13) illustrates the use of free *Ish* with modifying function and the bodily actions it is accompanied by.

### (13) *Ish* breathing

- 1 RS sure; right; [ so we' re breathing-  
 KR yeah[ I- I find that (when you are) -  
 KR\_gaze ■to camera----->
- 2 RS we' re breathing NOW; [ we' re breathing NOW;  
 KR [ I' m- (wo) -  
 KR\_gaze ----->  
 KR\_head +nods----->
- 3 RS [ but this is a different kind of BREATHing.  
 KR [ yeah.  
 KR\_gaze ---■down-----■up-----■  
 KR\_eyes \*closed-----\*  
 KR\_head ---- - +  
 KR\_mouth ◇lip smack◇
- 4 KR #ish; (. )  
 KR\_gaze -----■to camera■  
 KR\_eyebrows ●raised●  
 KR\_eyes \*upper lid raised\*  
 KR\_head +tilt to left+  
 Fig #Fig. 17
- 5 KR ish; #  
 KR\_gaze ■elsewhere----->  
 KR\_head +tilt to right----->  
 Fig #Fig. 18
- 6 KR we' re #ISH breathing.  
 KR\_gaze ----->>  
 KR\_eyes \*slit----->>  
 KR\_head -----+ +nods----+  
 Fig #Fig. 19



Figure 17.



Figure 18.



Figure 19.

Example (13) has been taken from a video call involving three participants: two hosts of a morning talk show, KR and RS, and their guest, an expert on breathing techniques (not part of the extract). Before the extract begins, the breathing expert explains to her audience the characteristics of a good breathing technique and what impact this has on the immune system. The extract starts with both hosts showing their understanding of this explanation using a variety of response tokens ('sure', 'right', and 'yeah'). This is followed by a competition about next speaker allocation since both hosts start simultaneously. Finally, RS takes the turn and claims that they are 'breathing now', but contrasts this with what was said before, claiming that 'this is a different kind of breathing'. His cohort qualifies this claim further by reacting twice with free *Ish*, thus implicating that the breathing technique they are using is very different from the one that boosts the immune system. KR then makes this implicature more explicit by saying that they are 'ish breathing'. *Ish* here is also a free morpheme, which modifies a verb, and questions its appropriateness.

This extract contains three instances of free *Ish*, all of which are accompanied by a different set of kinesic features but all serving similar functions. The first instance of free *Ish* (see Block 4 in the transcript) is accompanied by gaze aversion, raised eyebrows with the upper eyelid raised, too, and a head tilt to the speaker's left side (see Fig. 17). During a micropause that follows, she briefly looks to the camera. On the second instance of *ish*, KR tilts her head to her right side and looks somewhere else,<sup>4</sup> whereas her facial muscles rest in a neutral position (see Fig. 18). Finally, when she makes her implicature more explicit, she still avoids looking at the camera (but looks in a different direction), has her head slightly tilted, and narrows her eyes to a slit (see Fig. 19). Fig. 19 also shows that, in comparison with Fig. 18, her head is already slightly raised. This is because she prepares another head movement, that is, a nod, which begins on 'breathing' right after *ish*.

Gaze aversion might serve multiple purposes here. For one thing, KR might want to secure her turn and gaze aversion is a useful tool to display this intention both in face-to-face and video call interactions (Zima, *accepted*; Zima et al., 2019). Because of the overlapping talk that occurred before, she might want to make sure that it is her turn now. On the other hand, she might want to weigh the options, that is, whether

<sup>4</sup>This time, it seems as if she is looking at something, like a second monitor with notes or instructions, but given the methodology used in this article, we are unable to prove this. On the first instance of *ish*, though, she looks in an upward direction, making explicit gaze aversion more likely.

the use of the term *breathing* is appropriate in this context in the light of the explanation the breathing expert gave before. This function of explicit gaze aversion supports the function of approximate *ish* and, thus, enhances the effect. Since KR is the host of this show and its main purpose is to entertain their audience, this function is not unlikely. However, it needs to be mentioned again that gaze did not prove to be significant in the given study. In other words, while gaze aversion as a resource seems plausible for this example, it does not seem to be a systematic one.

As for facial movements, raising the upper eyelids, though this was not significant in the model, proved to be at least associated with free uses of *ish* and is illustrated in its first use in Example (13). Raising the upper eyelids (often in combination with raising the eyebrows) has been described to indicate the perception of something new and possible uncertainties surrounding this (Scherer, 1984; Smith, 1989). In Example (13), it might be that KR raises her eyebrows and upper eyelids to highlight the fact that – given the new input from an expert in breathing techniques – the way they breathe is not ‘proper’. On a more general level, modifying *ish* might be often accompanied by raised upper eyelids because it highlights some new aspect of the previous utterance, namely its questionable appropriateness in the given context and, hence, the need to modify the utterance in some respect without being explicit about it.

The reactions to the uses of free *Ish* are not shown in the extract above but can be briefly summarized. RS, the cohost, echoes the ‘ish’ and laughs about KR’s modification. KR then starts smiling with her cheeks raised. Smiling with cheeks raised also proved to be associated with free uses of *ish* (see Table 5 and Fig. 3) and indicates genuine positivity (the so-called ‘Duchenne smile’ as opposed to the ‘non-Duchenne smile’, which lacks the cheek raiser and can be perceived as insincere (Gunnery & Hall, 2015). In Example (13), the smile occurs significantly after uttering *ish* and, therefore, was not considered for the annotation. However, in other examples, the smile and the cheek raiser co-occurred with free *Ish*. Both the quantitative results reported above and the reaction to free *Ish* in Example (13) show its humorous potential. Speakers who smile and raise their cheeks while uttering free *Ish* seem to be aware of this potential and display this understanding.

Finally, the head movements performed in Example (13) need systematic attention. The first two instances of free *Ish* are accompanied by head tilts. As mentioned above, head tilts are often used to indicate disaffiliation with a third party (Debras, 2017; Debras & Cienki, 2012). In Example (13), though, it seems that disaffiliation is expressed not with a third party but with another interactant, here RS. In any case, the disaffiliating function of head tilts supports the function of free *Ish*. Since free *Ish* is used to modify a proposition that was uttered immediately before, the speaker of free *Ish* can display their distancing from this utterance with the help of the head tilt. Considering the association between free *Ish* and head tilts reported in Section 4, this seems to be a systematic, functional relationship. Another interesting observation in Example (13) is that KR eventually nods right after having uttered the third instance of *Ish*. Thus, after having considered whether their pulmonic actions can rightfully be called ‘breathing’ from some distance, she seems to come to conclude that her co-host is right and affirms his claim that they are ‘breathing now’. While this is not the only instance where free *Ish* is accompanied by (slow) head nods, this kind of pattern does not occur frequently enough to reach a statistically significant level.

## 8. Summary and conclusions

The present study shows that the different *ish* constructions are each a multimodal construction because, formally, they differ regarding their prosodic aspects of delivery to such an extent that they were significant in the generalized mixed-effects model that was fitted. In other words, their difference in duration, mean pitch height, and pitch variability as well as their level of prosodic integration are sufficiently large to distinguish them. Consequently, features like these need to be part of the usage-based constructional scheme. Interestingly, all of these features were prosodic ones for *ish* constructions. The limits of this article do not allow a full discussion of why only the prosodic features turned out to be significant in the model. One reason might be the interaction types that were considered in the study. Even though televised discourse presupposes their audience to attend to the visual stimuli as well, the acoustic channel is more reliable when the audience is inattentive or distracted. TV personalities other than actors might be aware of this and act accordingly. In any case, future studies working with spontaneous talk-in-interaction might be revealing in this respect.

Even though the kinesic features in this study did not reach a significant level in the model, some of them proved to be associated with the functions of the *ish* constructions, at least. These included the wiggle gesture and head movements as well as movements in the eye and mouth regions of the face. All of these features have meanings independent from the *ish* construction they occur with: the wiggle (performed by either hand, head, or both) is used to display inappropriateness, head tilts display disaffiliation, raised eyelids novelty, and smiles with raised cheeks humorlessness. Given that, it seems plausible to consider these independent, nonverbal form-function pairings, that is, constructions, which are associated with the *ish* construction by their function. When combined with a cross-modal collocation (Uhrig, 2018, 2019), they can interact with one another to arrive at cross-modal pragmatic constructs as illustrated in the examples above.

Another, albeit noncentral, result is the essential role entrenchment seems to play. Notwithstanding the fact that entrenchment was very crudely operationalized in this study as either having an entry in the OED or not, the study still gives some clues about its role for the use of nonverbal features. The results of the study suggest that lexicalized uses of constructions that are, from a usage-based perspective, daughter constructions are less often accompanied by nonverbal features than the constructs that are based on the schematic mother construction. This could be because they require more cognitive effort in production and reception (since the construct's meaning is not readily available) and the supportive function of the nonverbal features is a welcome asset. This is slightly different for free *Ish*, which is, technically, entrenched, but still supported by nonverbal features. This might be because the inferences free *Ish* triggers are context-dependent and features such as head tilts and/or smiles help the hearer in contextualizing the *Ish*. Likewise, due to its colloquial nature, free *Ish* might not be entrenched by all speakers to the same extent. Users might be aware of this and, therefore, might opt for adding nonverbal assets.

In sum, the present study could show that *ish* constructions can formally be distinguished by prosodic features matching meaning differences and are frequently accompanied by supporting nonverbal features. In this sense, they form a network of multimodal-*ish* constructions.

**Competing interest.** The authors declare none.

## References

- Barth-Weingarten, D. (2016). *Intonation units revisited: Cesuras in talk-in-interaction*. Benjamins.
- Bateman, J. (2016). Methodological and theoretical issues in multimodality. In N.-M. Klug & H. Stöckl (Eds.), *Handbuch Sprache im multimodalen Kontext* (pp. 36–74). De Gruyter.
- Bateman, J., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis – A problem-oriented introduction*. De Gruyter Mouton.
- Bochnak, M. R., & Csapak, E. (2014). A new metalinguistic degree morpheme. *Proceedings of Semantics and Linguistic Theory*, 24, 432–452.
- Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer (Version 6.0.52) [Windows]. Amsterdam.
- Bressem, J., & Müller, C. (2017). The ‘negative-assessment-construction’ – A multimodal pattern based on a recurrent gesture? *Linguistics Vanguard*, 3(s1), 1592–1604.
- Cánovas, C. P., & Valenzuela, J. (2017). Timelines and multimodal constructions: Facing new challenges. *Linguistics Vanguard*, 3(s1), 20161006.
- Debras, C. (2017). The shrug: Forms and meanings of a compound enactment. *Gesture*, 16(1), 1–34. <https://doi.org/10.1075/gest.16.1.01deb>
- Debras, C., & Cienki, A. (2012). Some uses of head tilts and shoulder shrugs during human interaction, and their relation to stancetaking [Conference presentation]. In *International conference on privacy, security, risk and trust and international conference on social computing*. Amsterdam, Netherlands.
- Eitelmann, M., Haugland, K., & Haumann, D. (2020). From engl-isc to whatever-ish: A corpus-based investigation of -ish derivation in the history of English. *English Language and Linguistics*, 24(4), 801–831. <https://doi.org/10.1017/S13606743190003>
- Ekman, P., & Friesen, W. V. (1978). *The facial action coding system: A technique for the measurement of facial movements*. Consulting Psychologists Press.
- ELAN. (2021). ELAN [Computer software] (version 6.2). Max Planck Institute for Psycholinguistics, The Language Archive. <https://archive.mpi.nl/ta/elan>
- Elff, M. (2022). Package ‘mclgfit’. Multinomial logit models, with or without random effects or overdispersion (version 0.9.4.2). <http://mclgfit.elff.eu/>
- Elvira-García, W. (2019). Two constructions, one syntactic form: Perceptual prosodic differences between elliptical and independent clauses in Spanish. In K. Beijering, G. Kaltenböck, & M. S. Sansiñena (Eds.), *Insubordination. Theoretical and empirical issues* (pp. 240–264). De Gruyter Mouton.
- Feyaerts, K., Brône, G., & Oben, B. (2017). Multimodality in Interaction. In B. Dancygier (Ed.), *The Cambridge handbook of cognitive linguistics* (pp. 135–156). Cambridge University Press.
- Goldberg, A. E. (2006). *Constructions at work. The nature of generalizations in language*. Oxford University Press.
- Gras, P., & Elvira-García, W. (2021). The role of intonation in construction grammar: On prosodic constructions. *Journal of Pragmatics*, 180, 232–247.
- Gunnery, S. D., & Hall, J. A. (2015). The expression and perception of the Duchenne smile. In A. Kostić & D. Chadee (Eds.), *The social psychology of nonverbal communication* (pp. 114–133). Palgrave Macmillan.
- Harris, T. (2020). Vagueness, context-sensitivity and scale structure of four types of adjectives with the suffix -ish. In R. Gergel & J. Watkins (Eds.), *Quantification and scales in change* (pp. 67–84). Language Science Press.
- Harris, T. (2021). *-ish/Ish: Aspects of a suffix turned free morpheme* [Doctoral dissertation, Universität Mannheim]. <https://madoc.bib.uni-mannheim.de/59992/>
- Holland, M. K., & Tarlow, G. (2012). Blinking and mental load. *Psychological Reports*, 31(1), 119–127. <https://doi.org/10.2466/pr0.1972.31.1.119>
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- Lüdtke, D. (2021). sjPlot: Data visualization for statistics in social science (Version R package version 2.8.10). <https://CRAN.R-project.org/package=sjPlot>
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–748.
- McNeill, D. (1992). *Hand and mind*. In *What gestures reveal about thought*. University of Chicago Press.
- McNeill, D. (2005). *Gesture and thought*. University of Chicago Press.

- Mittelberg, I. (2017). Multimodal existential constructions in German: Manual actions of giving as experiential substrate for grammatical and gestural patterns. *Linguistics Vanguard*, 3(s1), 20160047.
- Mondada, L. (2018). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106.
- Norde, M. (2009). *Degrammaticalization*. Oxford University Press.
- Norde, M. (2012). Degrammaticalization: Three common controversies. In K. Davids, T. Breban, L. Brems, & T. Mortelmans (Eds.), *Grammaticalization and language change: New reflections* (pp. 73–110). Benjamins.
- Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica*, 40(1), 1–18.
- Oltra-Massuet, I. (2017). Towards a morphosyntactic analysis of *-ish*. *Word Structure*, 10(1), 54–78.
- Oxford English Dictionary Online* (OED). (n.d.). *Oxford English Dictionary online*. Oxford University Press.
- Oxford Languages. (n.d.-a). *-ish, suffix*<sup>1</sup>. In *Oxford English Dictionary Online*. <http://www.oed.com/view/Entry/99965> (September 21, 2022).
- Oxford Languages. (n.d.-b). *-ish, adv*. In *Oxford English Dictionary Online*. <http://www.oed.com/view/Entry/260292> (September 21, 2022).
- Oxford Languages. (n.d.-c). *selfish, adj. and n*. In *Oxford English Dictionary Online*. <http://www.oed.com/view/Entry/175306> (October 6, 2022).
- Oxford Languages. (n.d.-d). *Pollyannaish, adj*. *Oxford English dictionary online*. <http://www.oed.com/view/Entry/263983> (October 6, 2022).
- Pentrel, M. (2013). Usage, status, and function of stand-alone *Ish*. In A. Ammermann, A. Brock, J. Pflaeging, & P. Schildhauer (Eds.), *Facets of linguistics. Proceedings of the 14th Norddeutsches Linguistisches Kolloquium 2013 in Halle (Saale)* (pp. 47–59). Peter Lang.
- Perniss, P. (2018). Why we should study multimodal language. *Frontiers in Psychology*, 9, 1109. <https://doi.org/10.3389/fpsyg.2018.01109>.
- Pöldvere, N., & Paradis, C. (2020). ‘What and then a little robot brings it to you?’ The reactive what-x construction in spoken dialogue. *English Language & Linguistics*, 24(2), 307–332.
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ravelle, W. (2022). *psych: Procedures for psychological, psychometric, and personality research (Version R package version 2.2.5)*. Northwestern University. <https://CRAN.R-project.org/package=psych>
- Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293–318). Psychology Press.
- Schoonjans, S. (2018). *Modalpartikeln als multimodale Konstruktionen: Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*. De Gruyter.
- Siefkes, M. (2015). How semiotic modes work together in multimodal texts: Defining and representing intermodal relations. *10plus1: Living Linguistics*, 1(1), 113–131.
- Smith, C. A. (1989). Dimensions of appraisal and physiological response in emotion. *Journal of Personality and Social Psychology*, 56(3), 339–353.
- Steen, F., & Turner, M. B. (2013). Multimodal construction grammar. In M. Borkent, B. Dancygier, & J. Hinnell (Eds.), *Language and the creative mind* (pp. 255–274). CSLI Publications.
- Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford University Press.
- Trousdale, G. (2011). *Ish*. Paper presented at ISLE 2, Boston, June 2011.
- Uhrig, P. (2018). *Multimodal constructions or crossmodal collostructions?* Paper presented at the international conference on construction grammar(s), Paris.
- Uhrig, P. (2019). *Crossmodal collostructions – A usage-based approach to multimodal communication in a construction grammar framework*. Paper presented at the 52nd annual meeting of the Societas Linguistica Europaea, Leipzig.
- Uhrig, P. (Ed.) (2020). *Multimodality in language and communication*. de Gruyter.
- Uhrig, P. (2021). *Multimodal corpus linguistics and machine learning methods for multimodal communication research*. Paper presented at the 2021 international conference on multimodal communication, Changsha.
- Vigliocco, G., Perniss, P., & Vinson, D. (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, 20130292. <https://doi.org/10.1098/rstb.2013.0292>
- Ward, N. G. (2019). *Prosodic patterns in English conversation*. Cambridge University Press.

- Wells, J. C. (2006). *English intonation: An introduction*. Cambridge University Press.
- Zima, E. (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, E. (accepted). Multimodale Analysen zu Fremd- und Selbstwahl in ZOOM-Videokonferenzen. In S. Kabatnik, L. Bülow, M.-L. Merten, & R. Mroczynski (Eds.), *Pragmatik multimodal*. Narr.
- Zima, E., & Bergs, A. (2017). Towards a multimodal construction grammar. *Linguistics Vanguard*, 3(s1), 20161006.
- Zima, E., Weiß, C., & Brône, G. (2019). Gaze and overlap resolution in triadic interactions. *Journal of Pragmatics*, 140, 49–69.

---

**Cite this article:** Lehmann, C. & Pentrel, M. (2024). Multimodal-ish: prosodic and kinesic aspects of bounded and free uses of *ish* *Language and Cognition* 16: 1–31. <https://doi.org/10.1017/langcog.2023.13>