



ARTICLE

“Talking heads” in Portuguese sign and spoken languages

Marisa Cruz  and Sónia Frota 

Center of Linguistics, University of Lisbon, Lisbon, Portugal

Corresponding author: Marisa Cruz; Email: marisac@edu.ulisboa.pt

(Received 15 November 2022; Revised 26 July 2024; Accepted 13 September 2024)

Abstract

Head and eyebrow movements have been reported as question markers in both spoken (e.g. Swerts & Kraemer, 2004) and sign languages (e.g., Zeshan, 2004). However, the relative weight of these visual cues in conveying prosodic meaning remains unexplored. This study examines, through a kinematic analysis, if (and how) the amplitude of head falling movements varies in statements *versus* questions, both in Portuguese Sign Language (LGP) and in the spoken modality of European Portuguese. The results show that the head falling movement plays a key role in conveying interrogativity in Portuguese, in varying degrees. In LGP, the head amplitude is larger than in the spoken modality, and the shape of the head movement varies across sentence types, thus showing the primary role of this visual cue in LGP prosodic grammar. In spoken Portuguese, although the head amplitude also differs between sentence types, the shape of the movement over time is always the same (falling), thus pointing to a secondary/complementary role in spoken Portuguese.

These findings not only contribute to the knowledge of the prosodic grammar of spoken and sign languages, but also challenge traditional language processing models, mostly focused on verbal language.

Keywords: Head movements; interrogativity; kinematics; signed and spoken modalities; Portuguese

1. Introduction

Several studies have shown that human communication is multimodal, involving nonverbally conveyed messages, referred to as visual cues, together with verbally conveyed messages. It has also been demonstrated that these co-speech visual cues are grammaticalized, expressing several semantico-pragmatic functions (Gibbon, 2009; Zellers, House & Alexanderson, 2016, *inter alia*). Moreover, visual cues and prosody are intertwined systems in multimodal communication, with a relevant role for sociopragmatics (Brown & Prieto, 2021). This also applies to signed languages: nonmanuals (facial expressions, head movements and body postures), in American Sign Language (ASL) and Israeli Sign Language, are grammaticalized (Sandler, 2005)



and are part of the intonational system of these signed languages (Dachkovsky & Sandler, 2009). However, the relative weight of visual cues in conveying prosodic meaning, in comparison with other cues, remains unexplored. The goal of the current study is to examine the relative weight of visual cues, focusing on nonmanuals (in sign language) and co-speech visual cues (in the spoken modality) produced to convey interrogativity in Portuguese.

By comparing interrogatives of thirty-five geographically and genetically different signed languages around the world, Zeshan (2004) observed that the range of crosslinguistic variation is large for some parameters, such as the structure of question-word paradigms. However, interestingly, the use of nonmanuals in questions exhibits more similarities (or less variability) across signed languages. Namely, head movements, together with eyebrow movements, have been reported to play a prosodic role as question markers. These nonmanuals were also shown to vary in form across question types, with yes-no questions being typically produced with eyebrow raising and head forward and down, and wh-questions being more frequently produced with eyebrow lowering combined with head forward or raising (Dachkovsky & Sandler, 2009; Dachkovsky, Healy & Sandler, 2013; Pfau & Quer, 2010; Sandler, 2005; Zeshan, 2004). In contrast with questions, statements do not seem to be marked by nonmanuals, as reported for ASL (Valli, Lucas, Mulrooney & Villanueva, 2011).

Differently from most signed languages, in Portuguese Sign Language (LGP) questions are mainly produced with eyebrow lowering (63% - Cruz, Swerts & Frota, 2019), and only the head movement changes between yes-no questions (head falling) and wh-questions (head raising). Statements, as observed for other signed languages, are predominantly not produced with any specific nonmanual marker (50% - Cruz, Swerts & Frota, 2019). When present, the most frequent nonmanual in statements is the head falling movement (35% - Cruz, Swerts & Frota, 2019). Despite the differences between LGP and other signed languages, a similar combination of nonmanuals was found, thus showing that the underlying system is similar: the same nonmanuals are used in a combinatory way (Cruz, Swerts & Frota, 2019). Although the relative weight of these nonmanuals in conveying prosodic meaning is largely unexplored, Cruz, Swerts and Frota (2019) advanced the proposal of considering eyebrow lowering as a question marker in LGP, because questions are mostly signaled by this nonmanual, and the head movement as a marker of the question type, because it varies across question types. However, it is important to mention that the authors also reported the production of some yes-no questions with the head falling movement as a single nonmanual, thus leading to a hypothetical ambiguity with the statements produced with the same nonmanual. These productions are important test cases to explore the relative weight of nonmanuals in conveying prosodic meaning.

In the spoken modality, several studies have shown that co-speech visual cues are organized into a system sharing several features with the prosody of spoken language (Krahmer & Swerts, 2009; Loehr, 2012; Mol, Krahmer, Maes & Swerts, 2012; Swerts & Krahmer, 2008). Eyebrows and head movements, as well as hands, are the most studied co-speech visual cues from a prosodic point of view. They have been shown to play a complementary role in the spoken modality, although the strength of co-speech visual cues is weaker than auditory cues for the distinction across sentence types (Crespo-Sendra, Kaland, Swerts & Prieto, 2013; Granström & House, 2004; House, 2002). Eyebrow movements have been described as a question marker (Cavé

et al., 1996; Purson et al., 1999), also playing a relevant role for the perception of focus (Krahmer, Ruttkay, Swerts & Wesselink, 2002), prominence (Krahmer & Swerts, 2007; Swerts & Krahmer, 2004, 2006, 2008), and for the distinction across sentence types and pragmatic meanings (Borràs-Comes & Prieto, 2011; Borràs-Comes, Kaland, Prieto & Swerts, 2014; Crespo-Sendra, Kaland, Swerts & Prieto, 2013). Head movements, like eyebrows, may vary in form and function (Kousidis, Malisz, Wagner & Schlangen, 2013). They have been considered as modality markers for uncertainty, they signal discourse structure and lexical repairs, and they may function as deictics or as backchannel giving/request (McClave, 2000; Poggi, D'Errico & Vincze, 2010). Although reported to pattern 'in predictable ways' (McClave, 2000: 856), it has been challenging to identify kinematic parameters that allow a systematic description of head gesture patterns, due to their dynamicity, variability and multidimensionality (Wagner, Malisz & Kopp, 2014).

In the spoken modality of European Portuguese (EP), eyebrow raising has been shown to be the dominant visual cue for questions across varieties, and thus a question marker, together with the nuclear pitch accent (Cruz, Swerts & Frota, 2015). In combination with eyebrow raising, the head falling movement also occurs in yes-no questions (56% in the standard variety – Cruz, Swerts & Frota, 2015). Because the visual cue characterizing the production of statements is the head falling movement alone (67%), Cruz, Swerts and Frota (2015) hypothesized that speakers could be sensitive to visual information to identify sentence types across varieties, especially in the absence of tonal contrast or in the presence of audiovisual mismatches. However, this hypothesis was not confirmed, as Cruz, Swerts and Frota (2017) have shown that intonational cues are more relevant than visual cues to identify sentence types. Nevertheless, the authors also concluded that visual cues might play a role in structural or linguistic marking, as participants' reaction times were lower when exposed to the audiovisual condition, comparatively to the audio only condition. This also suggests that processing two types of cues – auditory and visual – at the same time is not necessarily synonymous of a heavier cognitive process. Actually, visual cues might reduce the cognitive load (Goldin-Meadow, Nusbaum, Kelly & Wagner, 2001). This has been recently confirmed by Holler, Kendrick and Levinson (2018) and Holler and Levinson (2019), who showed that responses to questions accompanied by manual and/or head movements are faster than responses to questions conveyed auditorily only. Thus, in line with the proposal of an integrated system in language comprehension (Kelly, Özyürek & Maris, 2010), visual cues induce facilitation in processing, thus enhancing comprehension (Kita & Emmorey, 2023), and thinking (Kita, Alibali & Chu, 2017), besides speaking (see Clough & Duff, 2020, for a review on the role of gesture in communication and cognition).

It seems thus clear that visual cues are crucial for communication purposes, playing similar functions in different language modalities, although in different forms or degrees. This might be the case for eyebrow and head movements, which are both relevant for conveying sentence type and pragmatic meaning in LGP and in the spoken modality of Portuguese. As previously described, eyebrows are present in the production of yes-no questions, assuming different forms in both modalities (eyebrow lowering in LGP; eyebrow raising in the spoken modality). Head movements are also present in yes-no questions, assuming the same form in both modalities (head falling movement), and may be present in the production of statements (also head falling in both modalities).

2. Research questions and hypotheses

Since in LGP yes-no questions may be, although less frequently, produced with the head falling movement only (i.e., the eyebrow movement, typically considered as a question marker, is absent), a natural context arises for a hypothetical ambiguity with the production of statements characterized by the occurrence of the same nonmanual cue. However, communication for deaf signers is not disrupted as a result of this apparent overlap. We thus hypothesize that the amplitude of the head falling movement might be different between the two sentence types. If confirmed, this would mean that, in the absence of the eyebrow movement, the head assumes the role of question marker, thus reinforcing the dynamicity and multidimensionality of the head gesture (as suggested by Wagner, Malisz & Kopp, 2014). Additionally, it would show that, in signed languages, the interplay of nonmanuals is a complex and dynamic mechanism, subject to a relative weighting across nonmanuals used in combination.

Like in LGP, in the spoken modality of EP the head falling movement also occurs in statements and yes-no questions as single visual cue (17% - Cruz, Swerts & Frota, 2015). However, differently from LGP, in the spoken modality this does not lead to a hypothetical ambiguity, as the auditory (intonational) cues are not the same: yes-no questions are produced with a falling-rising melody, whereas statements are characterized by a falling melody. For this reason, and also taking into consideration the fact that speakers do not seem to rely on visual cues to identify sentence types, we do not expect to find a different amplitude of head movements between sentence types. If confirmed, this would allow us to conclude that the visual cues used to convey prosodic meanings are grammaticalized differently in each modality, what might impact the communication between signers and speakers.

To sum up, we hypothesize that, in LGP, the amplitude of the head falling movement might be larger in yes-no questions produced with the head falling movement only (i.e., without the eyebrow movement) than in statements, which are characterized by the occurrence of the same nonmanual. If confirmed, this will lead to the interpretation that the amplitude of the head falling movement is a question marker in LGP in the absence of eyebrow movement.

In the spoken modality of EP, we do not expect to find a different amplitude of head movement between the two sentence types (given that no ambiguity arises due to the presence of contrasting spoken intonational cues). The confirmation of this hypothesis will allow us to conclude that the visual cues used to convey prosodic meanings are grammaticalized differently in each modality.

In order to test these hypotheses, we conducted a kinematic analysis of the vertical displacement of head movements in statements and yes-no questions, both in LGP and in the spoken modality of Portuguese, following the methodological procedures detailed below.

3. Methods

3.1. Participants

Five native signers of LGP, all deaf women¹, aged between 20–50 years-old, participated in the study. Four of them reported congenital hearing deafness, while the fifth

¹Male deaf signers of LGP were also recorded. However, we only considered female deaf signers for the analysis to match the sample of the spoken modality, which follows the *InAPoP -Interactive Atlas of the*

reported acquired hearing deafness at the age of 6 months. They were videotaped in a quiet room of Associação Portuguesa de Surdos (APS, Portuguese Association of Deaf People) while performing an adapted version of the Discourse Completion Task (Félix-Brasdefer, 2009; Billmyer & Varghese, 2000) for LGP (Cruz, Martins & Frota, 2017). Participants interacted with an interpreter of LGP, also a linguist, who works on a daily basis in the deaf association where data collection took place. This ensured spontaneity in the task, as all participants were used to interacting with each other, besides the fact that they were in their natural environment. The sentences were elicited as follows: the interpreter of LGP presented several communicative contexts one by one, some of them supported by printed images, and asked the signers to produce what they would gesture in each communicative context as if they were there. A Canon – Legria video camera (model HF G25 AVCHD) was positioned in front of the participant, at a fixed distance across recordings, and the participant was sat in diagonal line with the interpreter. This setting was the same for all recordings.

Additionally, five native speakers of Standard European Portuguese (SEP), all women, aged between 20–45 years-old, were videotaped in a sound attenuated laboratory with a JVC video camera (model GY-HM11E), while performing the Discourse Completion Task, adapted for Portuguese within the project *InAPoP - Interactive Atlas of the Prosody of Portuguese* (Frota, coord., 2012-2015).

The current analysis was based on a sample of the collected data, as detailed below.

3.2. Dataset

The sample for analysis covers neutral statements (i.e., broad focus statements uttered as all-new utterances) and neutral yes-no questions (i.e., information-seeking questions where the information asked is new to the speaker). For LGP, we considered both the questions produced with the two most frequent nonmanuals (head falling movement, together with eyebrow lowering – Figure 1), and the potentially ambiguous yes-no questions, i.e., those produced with a head falling movement only (Figure 2).

Some yes-no questions produced with the two visual cues may also show a head forward movement as the head moves downwards. We ran an informal pilot perceptive study involving native signers, who were asked to watch videos of yes-no questions with and without the head forward movement, and to choose one out of 4 possible answers: (i) the signer wants to know something, without any previous related knowledge, (ii) the signer wants to know something, with previous related knowledge, (iii) the signer already knows the answer to the question, but she wants to confirm, (iv) other situation – please specify and explain why. This pilot revealed that the head forward movement is not interpreted as a question marker or as relevant information from a pragmatic point of view. For this reason, we considered these productions as conveying two visual cues for interrogativity.

Both statements and questions are produced with a falling head movement in spoken Portuguese (Figure 3). However, the potential ambiguity between statements and questions with one visual cue that characterizes LGP does not occur in the spoken modality, as auditory cues are distinct between sentence types, namely the nuclear

Prosody of Portuguese project's criterion of collecting female voices only, in order to avoid artifacts in the acoustic signal, such as creaky voice.



Figure 1. Yes-no question produced with a falling head movement together with eyebrow lowering – two visual cues –, in LGP. Video file available in the supplementary materials, accessible at journals.cambridge.org.

contour is different, and were shown to be more relevant than visual cues to perceive questions (Cruz, Swerts & Frota, 2017).

The most frequent production of a yes-no question in the spoken modality is illustrated in Figure 4, with the falling-rising intonational cues accompanied with the visual cues - head falling movement together with eyebrow raising.

All these sentences were previously annotated in ELAN, version 6.0 (Wittenburg, Brugman, Russel, Klassmann & Sloetjes, 2006; Max Planck Institute for Psycholinguistics, The Language Archive, 2020), with detailed information on head and eyebrow movements time-aligned with intonational cues (for the spoken modality – see Cruz, Swerts & Frota 2015 for further details) or with manuals (for LGP – see Cruz, Swerts & Frota, 2019 for details). FACS – Facial Action Coding System (Ekman, Friesen & Hager, 2002) was used to annotate the visual cues, and the whole movement of each visual cue was taken as an event. After this annotation, done by a linguist trained in audiovisual annotation and in LGP, the utterances were selected for the kinematic analysis by filtering the sentence types and the labels of interest for the current study, which were related with the type and number of visual cues involved (e.g., yes-no questions with both visual cues or with the head movement alone). In total, 101 utterances were considered, with the following distribution: (i) 26 sentences from LGP, including 9 statements, 13 yes-no questions produced with both nonmanuals, and 4 yes-no questions produced with the head only; (ii) 75 sentences from the spoken modality, comprising 18 statements, with the common nuclear contour H + L* L% (Frota, 2002), 39 yes-no questions produced with the most frequent nuclear contour H + L* LH% (Frota, 2002) and both visual cues (head falling movement, together with eyebrow raising), and 18 yes-no questions with the same melodic pattern, but involving the head movement only.

All 101 utterances were labeled in ELAN by one additional annotator, naïf in audiovisual annotation and in LGP, who performed a blind annotation, i.e., he only had access to two tiers (one for the eyebrows and another one for the head movements) with blank intervals time-aligned with the video. For the LGP data, no information was provided about glosses, sentence type or meaning being



EU IR CINEMA
 I TO-GO CINEMA
 ('I'm going to the cinema.')

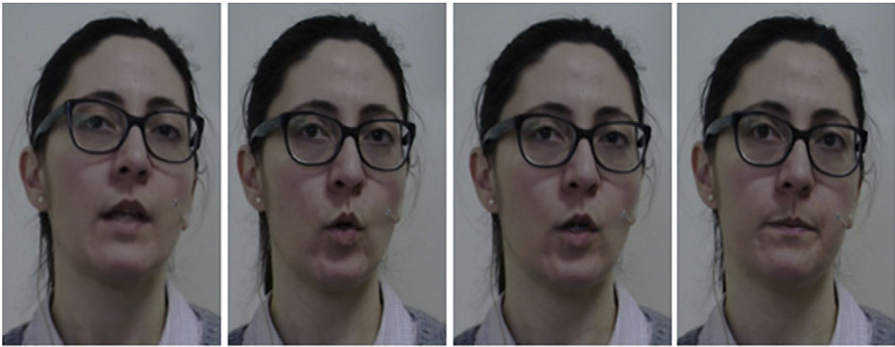


TU IR CINEMA
 YOU TO-GO CINEMA
 ('Are you going to the cinema?')

Figure 2. Statement produced with a falling head movement (top sequential frames), and yes-no question produced with a falling head movement only – one single visual cue – (bottom sequential frames), in LGP. Video files available in the supplementary materials, accessible at journals.cambridge.org.

produced. However, and in order to avoid biased annotations, a third tier was added to the annotator's grids in order to call his attention to nonmanuals conveying lexical information. For instance, the adverb ONTEM ('yesterday') is manually articulated with the pointing finger moving backwards above the shoulder and a head up movement, which plays a role in the lexical/morphological component of the grammar, thus unrelated to the prosody of statements and questions.

In order to control for variability of labels used, the ELAN grids included a controlled vocabulary containing a set of labels for the eyebrows (neutral, raising, lowering) and another set for the head (neutral, up-down, down-up, left-right, tilt left, tilt right, forward, backward). The naïf annotator only had to double-click in each time interval and to choose the label that better described the eyebrow and the head movement occurring in its time span. At the end, he was asked to revise all his



TEM COMPOTA.
 ('(He/She) has jam.')



TEM COMPOTA?
 ('Do you have jam?')

Figure 3. Statement (top sequential frames) and yes-no question (bottom sequential frames), both produced with a falling head movement in the spoken modality of Portuguese. Video files available in the supplementary materials, accessible at journals.cambridge.org.

annotations in order to ensure internal consistency in the labelling task performed, both within and across modalities.

Cohen's kappas were then computed, using SPSS (version 26) (IBM Corp., 2019), to generate pairwise inter-rater agreement scores. Following Landis and Koch's interpretation of kappa (Landis & Koch, 1977), an overall substantial agreement was found between the trained linguist annotator and the naïf annotator for the eyebrows (Cohen's $\kappa = .78$) and an overall slight agreement for the head (Cohen's $\kappa = .15$). When considering the modality annotated, we observed a moderate agreement for the eyebrows in LGP (Cohen's $\kappa = .52$) and an almost perfect agreement for the same visual cue in the spoken modality Cohen's $\kappa = .84$). For head movements, a slight agreement was found in LGP Cohen's $\kappa = .11$) and a slight agreement in SEP Cohen's $\kappa = .10$). This clearly shows that (i) annotating LGP, in general, is harder, which was expected, as the naïf annotator does not know LGP; (ii) annotating head movements is harder than eyebrows because the set of labels available in the controlled vocabulary for the head movements was bigger, covering



CHOVEU?
(‘Did it rain?’)

Figure 4. Yes-no question produced in the spoken modality of Portuguese, with a falling-rising melody, and a falling head movement together with eyebrow raising. Video file available in the supplementary materials, accessible at journals.cambridge.org.

all rotation possibilities in the y and x axis. The eyebrows, in contrast, are not physiological dynamic as the head is. For these reasons, we decided not to exclude the sentences with unmatching labels from the dataset.

3.3. Analysis procedures

All the 101 utterances were prepared for the extraction of the respective kinematic information according to the procedure detailed in section 3.3.1. The kinematic measurements were then subject to statistical analysis, as described in section 3.3.2.

3.3.1. Kinematics

Using *Kinovea* (version 0.9.4), a free 2D motion analysis software, the head vertical displacement was tracked along the time series (ms) and extracted in pixels (px). For each video file, we added a marker to the participant’s chin (Figure 5), and we inspected the video frame by frame to manually correct the marker location when needed (for instance, when manuals are partially superimposed with the head movement, which results in a tracking loss).

Since we are dealing with productions of varying length in number of syllables and words, we automatically extracted the time-normalized vertical displacement to control for duration differences across sentences. This allowed us to merge the kinematics data of each sentence type per modality and to observe/compare the head peaks along the time series across sentence types, also considering the type of visual (and auditory) cues present in the signal, as described in sections 4.1, 4.2, and 4.3.

3.3.2. Statistics

The statistical analysis performed was similar for both data samples – LGP and spoken modality of Portuguese. Namely, Generalized Additive Models (GAMs) were run for each sample (Wood, 2017), as they allow to compare the head vertical displacement extracted in pixels over time, by modelling non-linear dependencies



Figure 5. Kinovea marker in the participant's chin for the kinematics analysis along the time series. First frame of the production of a yes-no question by a signer of LGP.

of this dependent variable and sentence type as predictor (statement, yes-no question with two visual cues, and yes-no question with one visual cue), via smooth functions. For each modality, we ran a GAM (with fixed predictors only) and a Generalized Additive Mixed Model (GAMM), with participant as random factor. All models were initially fitted using *fREML* (fast restricted maximum likelihood estimation), which is the default smoothing parameter estimation method obtained in GAM. However, in order to prevent oversmoothing, model fits were also checked using the function *gam.check()*. Since the *k-index* was not smaller than 1 in any model, the default smoothing parameter ($k = 10$) was kept, thus, none smoothing adjustments were needed. Then, to compare the kinematics of the head movement between modalities over time, we ran another GAMM with Modality (spoken versus signed) as predictor, additionally to the Sentence type. Because the mixed models ran for each sample exhibited a better fit than the models without the random factor, participant was here included as random factor (and only a GAMM was run). Additionally, the interaction Sentence type*Modality was also modeled over time, as we were interested in observing whether the head vertical displacement is different across sentence types, depending on the modality.

The dependent variable corresponds to a total of 7500 measurements for the spoken modality (i.e., 100 measurements for each sentence, along the normalized time series into 1 s) and 2600 measurements for LGP.

All statistical procedures were run in R software (version 4.4.1, R Core Team 2024), using R packages *mgcv* (Wood, 2011, 2017) for modelling, and R package *itsadug* to plot the models results (van Rij et al., 2017).

4. Results

The results are presented first for LGP, and then for the spoken modality.

4.1. Portuguese Sign Language (LGP)

Figure 6 shows head vertical displacement in LGP along the normalized time series, merged per sentence type, and comparing statements and yes-no questions. We observe that statements (blue line) are less marked by head movement, ranging from 4.54 to -4.59 pxs, than yes-no questions (gray and orange lines), which display a wider degree of displacement.

Within yes-no questions, those produced with the head as the single visual cue show the widest amplitude, ranging from 8.45 to -17.47 pxs (orange line), whereas yes-no questions produced with head and eyebrow movements present a smaller amplitude, ranging from 3.04 to -15.86 pxs (gray line). This seems to suggest that the head amplitude plays a relevant role in conveying interrogativity, and that it varies with the amount of nonmanuals in the signal, along the lines of our initial hypothesis.

To confirm whether head amplitude plays such a relevant role, a GAM was run with head vertical displacement as dependent variable, and sentence type as predictor, with yes-no question with two visual cues as reference level – Model 1. We also ran Model 2 (GAMM), with participant as random factor. Model 2 showed a better fit than Model 1 ($fREML = 9489.7$ versus $fREML = 9703.5$), explaining 35.6% of the deviance (contra 23.2% in Model 1). For this reason, we report the results from Model 2, plotted in Figure 7. We observed that the head vertical displacement differs across sentence types (statement versus reference level: $\beta = 6.74$, $SE = 0.44$, $t = 15.44$, $p < .001$; yes-no question with one visual cue versus reference level: $\beta = -2.22$, $SE = 0.60$, $t = -3.71$, $p < .001$).

As clearly shown in Figure 7, the head falling movement as the single visual cue for yes-no questions (blue curve) is considerably more pronounced than the same visual cue in statements (green curve), showing that these utterances are not ambiguous in production. The vertical head movement that co-occurred with eyebrow lowering in yes-no questions (red curve) is also different from the head movement produced in

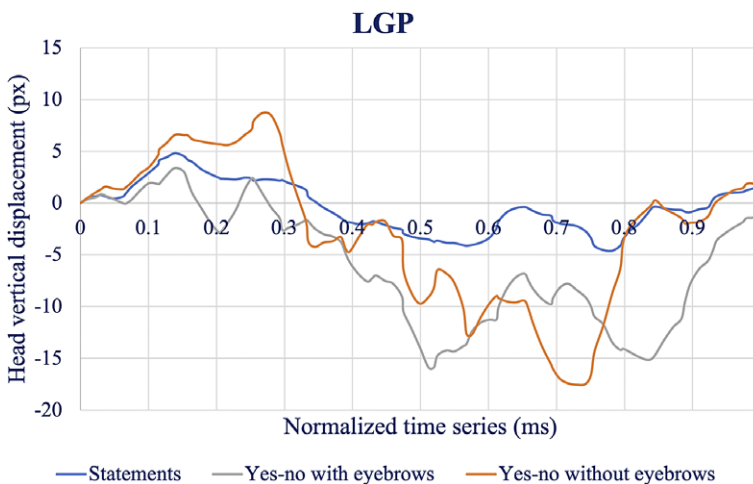


Figure 6. Head vertical displacement (pxs) in LGP, averaged per sentence type, along the normalized time series (ms). Statements are plotted in blue, yes-no questions produced with both cues (head and eyebrows) are plotted in gray, and yes-no questions produced with the head cue only (without eyebrows) are plotted in orange.

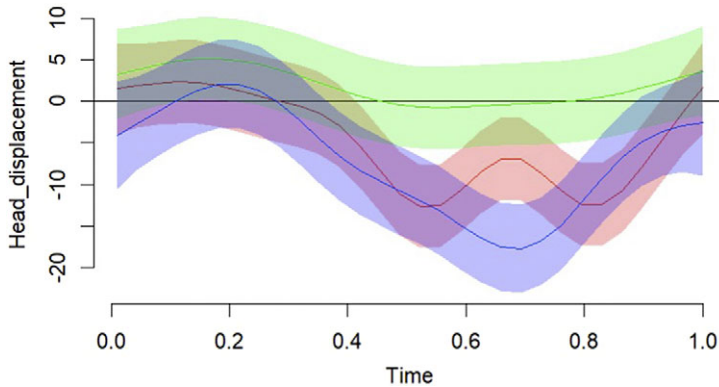


Figure 7. Non-linear smooths (fitted values) for statements (green), yes-no questions with two visual cues (eyebrows+head) (red), and yes-no questions with one visual cue (head) (blue) in LGP. 95% Confidence Intervals are shown by shaded bands.

statements (green curve), indicating that interrogativity is conveyed by a larger degree of head displacement. We may also observe in [Figure 7](#) that both curves start at closer height position, but the head in yes-no questions produced with eyebrow lowering moves downwards before 200 ms and attains a larger amplitude. When comparing the curves of both interrogatives (with two visual cues - red - and one visual cue - blue), two salient features emerge. First, when only the head is used, its falling movement is more pronounced than in yes-no questions produced with eyebrows, thus showing a larger amplitude. Second, there is also a difference at the beginning of the head movement: in yes-no questions with one visual cue, it starts at a lower position (in the vertical axis), which leads to a short rising movement (see the time-window of the first 200 ms), contrasting with the falling movement of the head in yes-no questions with two visual cues, which starts before 200 ms, as mentioned above.

Moreover, the shape of the curves is different across sentence types. Namely, statements clearly contrast with questions, as we may observe an initial peak, followed by a falling movement, and a final rising movement, corresponding to the retraction phase (McNeill, 1992). In questions, the shape of the movement is more complex and also differs depending on the amount of nonmanuals in the signal: yes-no questions produced with two visual cues (more complex) also present more complexity in the head movement – an initial peak followed by two (lower) peaks, reminding a binomial distribution², and then the retraction; yes-no questions with one visual cue (less complex) show a head movement with less complexity – an initial peak followed by a lower peak, and then the retraction.

These findings confirm that the head movement amplitude plays a relevant role in conveying interrogativity by distinguishing yes-no questions from statements, and that it varies with the amount of non-manuals in the signal.

²We are grateful to one of the reviewers, who called our attention to this aspect. The time alignment of the head movement peaks with the manuals will be inspected in future research.

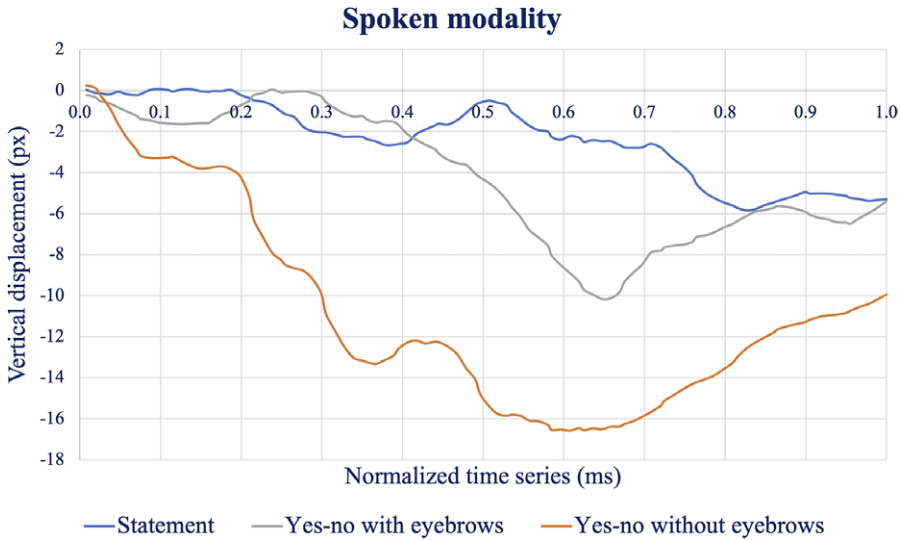


Figure 8. Head vertical displacement (pxs) in the spoken modality of Portuguese, averaged per sentence type, along the normalized time series (ms). Statements are plotted in blue, yes-no questions produced with both cues (head and eyebrows) are plotted in gray, and yes-no questions produced with the head cue only (without eyebrows) are plotted in orange.

Finally, we also observe that the highest displacement values of the head movement are attained at the same time-window (500–850 ms), independently of the sentence type. Given that nonmanuals are part of the intonational system of signed languages (Dachkovsky & Sandler, 2009), we may thus hypothesize that the head is a cue for the prosodic nucleus. However, this requires further analysis in order to check the exact time-alignment of the head movements with manuals, which is not possible in the current study, as sentences were time-normalized.

4.2. Spoken modality of Portuguese

Figure 8 shows head vertical displacement in the spoken modality along the normalized time series, per sentence type. We observe that statements (blue line) are less marked by the head movement, like in LGP, ranging from 0.07 to -5.85 pxs, than yes-no questions (gray and orange lines), which show a wider degree of head displacement.

Also like in LGP, in the spoken modality yes-no questions produced with the head as the single visual cue exhibit the widest amplitude, ranging from 0.24 to -16.59 pxs (orange line), whereas yes-no questions produced with head and eyebrow movements present a smaller amplitude, ranging from 0.04 to -10.16 pxs (gray line). Again, the head amplitude seems to play a relevant role in conveying interrogativity, especially when it is the single visual cue and even in the presence of different auditory cues, despite the fact that visual cues were previously found to be less relevant than auditory ones to perceive questions (Cruz, Swerts & Frota, 2017).

To confirm whether head amplitude plays such a relevant role, a GAM was run with head vertical displacement as dependent variable, and sentence type as predictor,

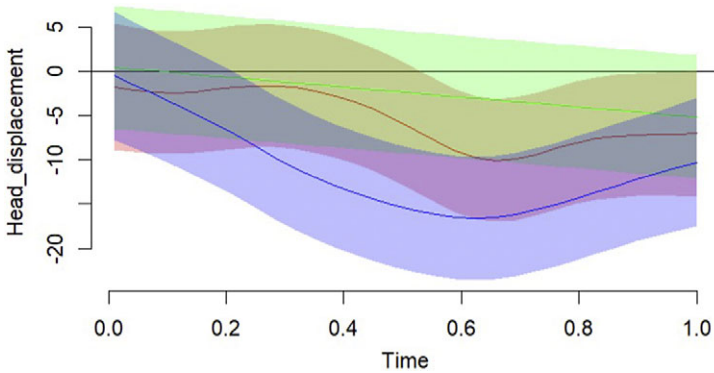


Figure 9. Non-linear smooths (fitted values) for statements (green), yes-no questions with two visual cues (eyebrows+head) (red), and yes-no questions with one visual cue (head) (blue) in the spoken modality. 95% Confidence Intervals are shown by shaded bands.

with yes-no question with two visual cues as reference level – Model 1. We also ran Model 2 (GAMM), with participant as random factor. Model 2 showed a better fit than Model 1 ($fREML = 29607$ versus $fREML = 30330$), explaining 26% of the deviance (contra 9.83% in Model 1). We thus report the results from Model 2, plotted in Figure 9. As in LGP, in the spoken modality, the head vertical displacement differs across sentence types (statement versus reference level: $\beta = 3.08$, $SE = 0.38$, $t = 8.17$, $p < .001$; yes-no question with one visual cue versus reference level: $\beta = -6.09$, $SE = 0.42$, $t = -14.6$, $p < .001$).

As in LGP, in the spoken modality, the head falling movement as single visual cue for yes-no questions (blue curve) is more pronounced than the same visual cue in statements (green curve), suggesting that these productions are not visually ambiguous, although the auditory cues already distinguished between the sentence types. In yes-no questions, against our hypothesis and in line with what is described in section 4.1 for LGP, the amplitude of the head falling movement is larger when one single visual cue is available (blue curve) than when both visual cues are present (red curve).

Differently from LGP, however, in the spoken modality all curves start at a closer height position and the head moves downwards in all sentence types. In yes-no questions with one visual cue, there is a clear difference between LGP and the spoken modality: in the latter, the falling head movement starts from the very beginning of the production, contrasting with the delayed falling movement in the same sentence type in LGP.

A comparison between the shape of the curves in the spoken modality with those observed in LGP (Figure 7) reveals that, unlike in LGP, in the spoken modality there is no shape variation, as all sentence types exhibit an initial peak, followed by a falling movement, and a final rising movement, corresponding to the retraction phase (McNeill, 1992). This leads us to hypothesize that, because visual information is less relevant than auditory information for speakers, it is less distinctive than in LGP. Nevertheless, the highest displacement values of the head movement in the spoken modality are attained at the same time-window as in LGP (500–850 ms), independently of the sentence type. In the spoken modality, this is probably time-aligned with the nuclear prosodic word. If so, this would be in line with our hypothesis that the

head, in LGP, is a cue for the prosodic nucleus. However, as mentioned above, this requires further analysis in order to check the exact time-alignment of the head movements with speech.

These findings lead us to conclude that the degree of head vertical displacement is similar in LGP and in the spoken modality, which goes against our initial hypothesis that the magnitude of the head movement would be smaller in the spoken modality, given the presence of contrasting auditory cues across sentence types. The difference across modalities seems to lie on the complexity of the head movement shape – more complex in LGP than in the spoken modality –, which seems to be related with the weight of visual cues in the signal. Namely, in the spoken modality visual cues are not as relevant as auditory cues to convey prosodic meaning, whereas they are critical in the signed modality.

4.3. LGP versus Spoken modality of Portuguese

Figure 10 depicts the kinematics of head movement per sentence type across modalities.

To examine the degree of similarity in the kinematics of the head movement between modalities (with both modalities showing a wider amplitude when the eyebrows are not present), we ran a GAMM with Modality (spoken versus signed) and sentence type as predictors, with spoken modality as reference level. Because the mixed models ran for each sample exhibited a better fit than the models without the random factor, participant was here included as random factor (Model 1). Then, the interaction Sentence type*Modality was also modeled over time, additionally to Modality and Sentence type as predictors (Model 2), as we were interested in observing whether the head vertical displacement is different across sentence types,

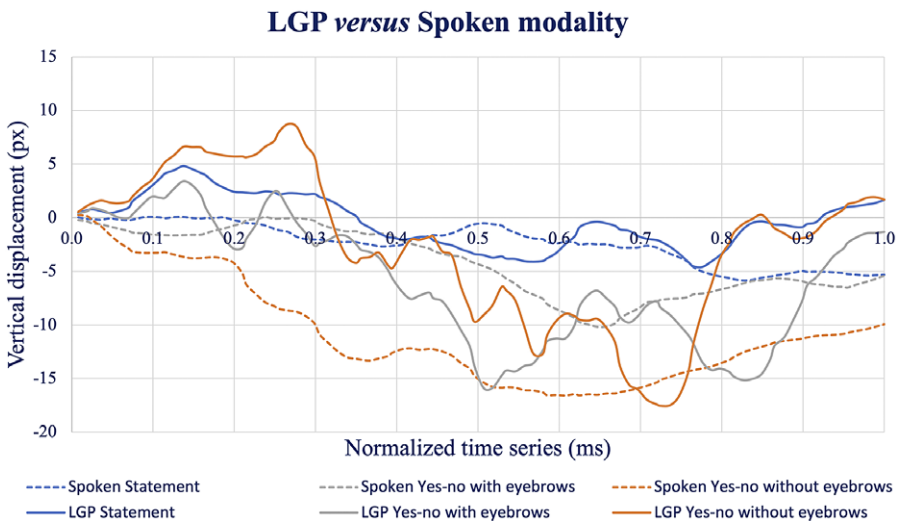


Figure 10. Head vertical displacement (pxs) in LGP (solid lines) and in the spoken modality of Portuguese (dotted lines), averaged per sentence type, along the normalized time series (ms). Statements are plotted in blue, yes-no questions produced with both cues (head and eyebrows) are plotted in gray, and yes-no questions produced with the head cue only (without eyebrows) are plotted in orange.

depending on the modality. Model 2, with the interaction term, showed a better fit than Model 1 ($fREML = 39913$ versus $fREML = 40051$), explaining 17.3% of the deviance (contra 15% in Model 1). We thus report the results of Model 2. We found that the head vertical displacement differs between modalities ($\beta = -2.03$, $SE = 0.42$, $t = -15.52$, $p < .001$), and that the interaction Sentence Type*Modality shows a significant effect. This means that the difference between the head vertical displacement in statements and yes-no questions with two visual cues is not the same across modalities ($\beta = 2.68$, $SE = 0.67$, $t = 4.00$, $p < .001$). Similarly, the difference between LGP and the spoken modality in the contrast between yes-no questions with one visual cue and two visual cues is not the same ($\beta = 14.00$, $SE = 0.84$, $t = 16.59$, $p < .001$).

As shown in Figure 10, the head amplitude is larger in LGP than in the spoken modality of Portuguese. This finding, together with the significant different kinematics per sentence type across modalities, strongly suggests that although both modalities make use of head movement to convey interrogativity in phonologically relevant ways for both grammars, head movement displays realizational differences, being more prominent in LGP.

5. Discussion

This study had the main goal of examining if (and how) the amplitude of head movements conveys interrogativity both in Portuguese Sign Language (LGP) and in the spoken modality of European Portuguese. In the light of findings from previous work (Cruz, Swerts & Frota, 2019), we hypothesized that the amplitude of the head falling movement in LGP would be larger in yes-no questions produced without eyebrow movement than in statements, thus functioning as a disambiguating cue, taking the role of question marker. Also in the light of findings from previous work (Cruz, Swerts & Frota, 2015, 2017), we hypothesized that in the spoken modality a different amplitude of head movements between statements and yes-no questions would not be observed, given the contrasting auditory (intonational) cues across sentence types.

In LGP, we found that the head falling movement as the single visual cue for yes-no questions was significantly more pronounced than the same visual cue in statements, thus showing that these utterances were not ambiguous in production, and that interrogativity seems indeed to be conveyed by a larger degree of head displacement. Importantly, this feature was not limited to potentially ambiguous sentences in LGP. When eyebrow lowering and the head movement co-occurred, we also found a significantly wider head amplitude than in statements. These findings suggest that the presence of eyebrow movement might not be as important as the head movement in conveying interrogativity in LGP, probably because eyebrow movement does not vary across question types in LGP, differently from the majority of signed languages (Dachkovsky & Sandler, 2009; Dachkovsky, Healy & Sandler, 2013; Pfau & Quer, 2010; Sandler, 2005; Zeshan, 2004). Moreover, a smaller head displacement was found in yes-no questions with eyebrow lowering than in yes-no questions without the eyebrow lowering cue. Therefore, our hypothesis that in the absence of the eyebrow cue the head movement would take the role of question marker seems to be confirmed, which reinforces the dynamicity and multidimensionality of the head gesture (Wagner, Malisz & Kopp, 2014), and shows that the

combination of nonmanuals is a complex and dynamic mechanism, subject to a relative weighting across nonmanuals.

This dynamicity of the head movement in LGP can also be observed in its shape over time, especially in yes-no questions produced with the eyebrow lowering cue. The more complex the nonmanuals, in terms of amount of cues, the more complex the head shape over time. This leads to the interpretation that the head plays a primary role in conveying interrogativity in LGP.

Similarly to LGP, in the spoken modality the head falling movement as single visual cue for yes-no questions was significantly more pronounced than the same visual cue in statements and in yes-no questions produced with two visual cues (eyebrow raising and head movement). This was unexpected due to previous findings from perception showing that speakers rely on the different intonational cues (falling pitch in statements; falling-rising pitch in yes-no questions) to identify sentence types (Cruz, Swerts & Frota, 2017). Thus, the hypothesis that head movement would not play a role in the spoken modality seems not to be confirmed. However, our findings related with the head movement shape over time, namely, the fact that it is always falling, independently of the sentence type, suggest that head movement is less relevant to convey interrogativity in the spoken modality than in LGP, indicating that the visual cues used to convey interrogativity might be grammaticalized differently across modalities. These different grammaticalization paths of visual cues in spoken and signed languages are, from our point of view, supported by the biology of each language modality: visual cues in signed languages are vital for communication and they are fully visible, thus being also faster recognized than words (Emmorey, 2023), whereas in the spoken modality this crucial role is played by vocal articulators and the related biological codes (Gussenhoven, 2002, 2004).

The current study has three important limitations. Our dataset was produced by a small sample of participants, and thus the potential generalization of the present findings to wider populations should be interpreted with caution. In addition, for the current study it was not possible to relate the semantic-pragmatic content of speech/signs, beyond sentence type, to the head movement trajectories, given the use of normalized time series. Finally, our option of analysing the head movement in the y axis only (because we were focused on its vertical displacement) can also be seen as a limitation, as the head is physically dynamic. Future research is needed to address these shortcomings.

6. Conclusion

By and large, the present findings seem to support the view that gestural features of language and prosody are interconnected or integrated in the marking of semantic and pragmatic meanings (Brown and Prieto, 2021). How these two communication vectors are time-aligned in order to embrace the holistic nature of sociopragmatics is left for future research avenues (Cruz & Frota, *in progress*). Also of relevance is the investigation of how head and eyebrow movements are time-coordinated with each other (Cruz & Frota, 2023), and whether, similarly to what we observed for the head movement, the amplitude of the eyebrow movement is semantico-pragmatically relevant.

Last but not least, and in line with Sandler (2005), the comparison of visual cues in both modalities allowed us to suggest that the two systems use the degree of head

movement to convey interrogativity, albeit with differences in its grammaticalization and realization. Even in the spoken modality, in which auditory cues are preferred over visual cues to identify sentence types, the head movement varies, thus pointing to the suggestion that this feature might be a cross-modality linguistic property, with a relevant role in interaction.

These findings also challenge traditional language processing models, mainly focused on verbal language. Although this was not the focus of the current study, in line with Kelly, Özyürek and Maris (2010), *inter alia*, we speculate that the amplitude of head movement might be relevant for a faster and accurate identification of sentence types, thus enhancing comprehension. Future studies should address this possibility. Importantly, how these cognitive processes apply and evolve in a communicative context involving both modalities remains an open question, as it is not yet known whether the processing of the interaction between language and gesture differs between spoken and signed languages (Kita & Emmorey, 2023).

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/langcog.2024.63>.

Acknowledgments. This research was supported by Fundação para a Ciência e a Tecnologia (UIDB/00214/2020). We thank all the signers and speakers for their collaboration, as well as *Associação Portuguesa de Surdos (APS)* and the Lisbon Baby Lab, where video recordings took place. Special thanks are also due to the anonymous reviewers for their precious comments and questions, as well as to Pedro Jordão for the blind annotations of the data, and Chao Zhou for helping with the statistical analysis.

Competing interest. The authors declare none.

References

- Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests. *Applied Linguistics*, 21(4), 517–552.
- Borràs-Comes, J., & Prieto, P. (2011). ‘Seeing tunes’. The role of visual gestures in tune interpretation. *Laboratory Phonology*, 2(2), 355–380.
- Borràs-Comes, J., Kaland, C., Prieto, P., & Swerts, M. (2014). Audiovisual correlates of interrogativity: A comparative analysis of Catalan and Dutch. *Journal of Nonverbal Behavior*, 38, 53–66.
- Brown, L., & Prieto, P. (2021). Gesture and prosody in multimodal communication. In M. Haugh, D. Z. Kadar & M. Terkourafi (Eds.), *The Cambridge Handbook of Sociopragmatics* (pp. 430–453). Cambridge University Press.
- Cavé, C., Guaïtella, I., Bertrand, R., Santi, S., Harlay, F., & Espesser, R. (1996). About the relationship between eyebrow movements and F0 variations. *Proceedings of the International Conference on Spoken Language Processing*, vol. 4, Philadelphia, 2175–2178.
- Clough, S., & Duff, M. (2020). The Role of Gesture in Communication and Cognition: Implications for Understanding and Treating Neurogenic Communication Disorders. *Frontiers in Human Neuroscience*, 14(323). doi: 10.3389/fnhum.2020.00323
- Crespo-Sendra, V., Kaland, C., Swerts, M., & Prieto, P. (2013). Perceiving incredulity: The role of intonation and facial gestures. *Journal of Pragmatics*, 47, 1–13. doi: 10.1016/j.pragma.2012.08.008
- Cruz, M., & Frota, S. (in progress). Temporal and/or prosodic coordination between head and eyebrow peaks in statements and yes-no questions? Evidence from spoken European Portuguese and Portuguese Sign Language.
- Cruz, M., & Frota, S. (2023). Gestures-Prosody (a)synchrony: Temporal interaction between head and eyebrow peaks in statements and yes-no questions. In R. Skarnitzl & J. Volin (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 4181–4185). Guarant International. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2023/ICPhS_2023_PROCEEDINGS.pdf

- Cruz, M., Martins, M., & Frota, S. (2017). *Discourse Completion Task – Adapted version for data collection in LGP*. Available at <http://labfon.lettras.ulisboa.pt/InAPoP/methodology.html>
- Cruz, M., Swerts, M., & Frota, S. (2015). Variation in tone and gesture within language. In The Scottish Consortium for ICPHS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, University of Glasgow, UK. ISBN 978-0-85261-941-4. Paper number 452 retrieved from <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0452.pdf>
- Cruz, M., Swerts, M., & Frota, S. (2017). The role of intonation and visual cues in the perception of sentence types: Evidence from European Portuguese varieties. Special collection *Multiplicity of cues and functions in prosody* (ed. by A. Leemann and B. Post). *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 8(1), 23. DOI: <http://doi.org/10.5334/labphon.110>
- Cruz, M., Swerts, M., & Frota, S. (2019). Do visual cues to interrogativity vary between language modalities? Evidence from spoken Portuguese and Portuguese Sign Language. *Proceedings of AVSP 2019 – International Conference on Auditory-Visual Speech Processing*, August 10–11, 2019, Melbourne, Australia. https://avsp2019.loria.fr/wp-content/uploads/2019/07/AVSP_2019_paper_12.pdf
- Dachkovsky, S., & Sandler, W. (2009). Visual intonation in the prosody of a sign language. *Language and Speech*, 52(2/3), 287–314.
- Dachkovsky, S., Healy, C., & Sandler, W. (2013). Visual intonation in two sign languages. *Phonology*, 30(2), 211–252. doi:10.1017/S0952675713000122
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System*. Salt Lake City, UT: A Human Face.
- ELAN – Linguistic Annotator (Version 6.0) [Computer software]. (2020). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. [<http://tla.mpi.nl/tools/tla-tools/elan/>].
- Emmorey, K. (2023). Ten things you should know about sign languages. *Current Directions in Psychological Science*, 32(5), 387–394. doi:10.1177/09637214231173071.
- Félix-Brasdefer, J. C. (2009). Data collection methods in speech act performance: DCTs, role plays, and verbal reports. In E. Usó Juan & A. Martínez-Flor (Eds.), *Speech act performance: Theoretical, Empirical, and Methodological Issues* (pp. 41–56). Amsterdam: John Benjamins Publishing.
- Frota, S. (2002). Nuclear falls and rises in European Portuguese: A phonological analysis of declarative and question intonation. *Probus* 14(1) (Special Issue on *Intonation in Romance*, ed. by J.-I. Hualde), 113–146. doi: 10.1515/prbs.2002.001
- Frota, S. (Coord.) (2012–2015). *Interactive Atlas of the Prosody of Portuguese Project* (PTDC/CLE-LIN/119787/2010), funded by FCT – Fundação para a Ciência e a Tecnologia, Portugal.
- Gibbon, D. (2009). Gesture theory is linguistics: On modelling multimodality as prosody. *Proceedings of the Twenty-third Pacific Asia Conference on Language, Information and Computation*, pp. 9–18.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, 12, 516–522. doi: 10.1111/1467-9280.00395
- Granström, B., & House, D. (2004). Audiovisual representation of prosody in expressive speech communication. In B. Bel & I. Marlin (Eds.), *Proceedings of the International Conference on Speech Prosody 2004*, 393–396.
- Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology. *Proceedings of Speech Prosody 2002* (pp. 47–57), Aix-en-Provence.
- Gussenhoven, C. (2004). Paralinguistics: Three Biological Codes. In *The Phonology of Tone and Intonation* (pp. 71–96). Cambridge University Press.
- Holler, J., & Levinson, S. C. (2019). Multimodal Language Processing in Human Communication. *Trends in Cognitive Sciences*, 23(8). doi:10.1016/j.tics.2019.05.006
- Holler, J., Kendrick, K.H., & Levinson, S.C. (2018). Processing language in face-to-face conversation: Questions with gestures get faster responses. *Psychonomic Bulletin & Review*, 25, 1900–1908. <https://doi.org/10.3758/s13423-017-1363-z>
- House, D. (2002). Intonation and visual cues in the perception of interrogative mode in Swedish. *Proceedings of the International Conference on Spoken Language Processing, 1957–1960*.
- Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two Sides of the Same Coin: Speech and Gesture Mutually Interact to Enhance Comprehension. *Psychological Science*, 21(2), 260–267. <https://doi.org/10.1177/0956797609357327>
- Kinovea 0.9.4. <https://www.kinovea.org/>

- Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, 124, 245–266. doi: 10.1037/rev0000059
- Kita, S., Emmorey, K. (2023). Gesture links language and cognition for spoken and signed languages. *Nature Reviews Psychology*, 2, 407–420. <https://doi.org/10.1038/s44159-023-00186-9>
- Kousidis, S., Malisz, Z., Wagner, P., & Schlangen, D. (2013). Exploring annotation of head gesture forms in spontaneous human interaction. In *Proceedings of the Tilburg Gesture Meeting (TiGeR 2013)*.
- Krahmer, E., Ruttkay, Z., Swerts, M., & Wesseling, W. (2002). Pitch, eyebrows and the perception of focus. In B. Bel (Ed.), *Proceedings of the Speech Prosody 2002*, 443–446.
- Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57(3), 396–414. doi: 10.1016/j.jml.2007.06.005
- Krahmer, E., & Swerts, M. (2009). Audiovisual prosody—introduction to the special issue. *Language and speech*, 52(2–3), 129–133.
- Landis, J., & Koch, G. (1977). Measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Loehr, D. P. (2012) Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology. Special Issue Gesture as Language, Gesture and Language* 2, 3(1), 71–89.
- McClave, E. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics* 32, 855–878.
- McNeill, D. (1992). Guide to Gesture Classification, Transcription, and Distribution. In *Hand in Mind. What Gestures Reveal about Thought* (pp. 75–104). The University of Chicago Press.
- Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2012). Adaptation in gesture: Converging hands or converging minds?. *Journal of Memory and Language*, 66(1), 249–264. <https://doi.org/10.1016/j.jml.2011.07.004>.
- Pfau, R., & Quer, J. (2010). Nonmanuals: Their grammatical and prosodic roles. In D. Brentari (Ed.), *Sign Languages (Cambridge Language Surveys)* (pp. 381–402). Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511712203.018
- Poggi, I., D’Errico, F., & Vincze, L. (2010). Types of Nods. The polysemy of a social signal. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*.
- Purson, A., Santi, S., Bertrand, R., Guaitella, I., Boyer, J., & Cavé, C. (1999). The relationships between voice and gesture: Eyebrow movements and questioning. *Proceedings of European Conference of Speech Communication and Technology* 6, Budapest, Hungary, 1735–1739.
- R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Sandler, W. (2005). An overview of sign language linguistics. In K. Brown (Ed.), *Encyclopedia of language and linguistics*, Vol 11, 2nd Edition (pp. 328–338). Oxford, UK: Elsevier.
- IBM Corp. Released 2019. *IBM SPSS Statistics for Macintosh, Version 26.0*. Armonk, NY: IBM Corp.
- Swerts, M., & Krahmer, E. (2004). Congruent and incongruent audiovisual cues to prominence. In B. Bel & I. Marlin (Eds.), *Proceedings of the International Conference on Speech Prosody 2004*, 69–72.
- Swerts, M., & Krahmer, E. (2006). The importance of different facial areas for signalling visual prominence. *Proceedings of the International Conference on Spoken Language Processing (Interspeech 2006)*. Pittsburgh, Pa, USA: ISCA.
- Swerts, M., & Krahmer, E. (2008). Facial expression and prosodic prominence: Effects of modality and facial area. *Journal of Phonetics*, 36(2), 219–238.
- Valli, C., Lucas, C., Mulrooney, K. J., & Villanueva, M. (2011) *Linguistics of American Sign Language. An Introduction*, 5th edition. Washington DC: Gallaudet University Press.
- van Rij, J., Wieling, M., Baayen, R. H., and van Rijn, H. (2017). itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. *R Package Version*, 2.
- Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209–232. doi: 10.1016/j.specom.2013.09.008
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *5th international conference on language resources and evaluation (LREC 2006)*, 1556–1559.
- Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 73, 3–36. doi: 10.1111/j.1467-9868.2010.00749.x

- Wood, S. N. (2017). *Generalized Additive Models: An Introduction With R* (2nd Ed.). Boca Raton, MA: CRC Press.
- Zellers, M., House, D., & Alexanderson, S. (2016). Prosody and hand gesture at turn boundaries in Swedish. *Proceedings of the International Conference on Speech Prosody*, pp. 831–5.
- Zeshan, U. (2004). Interrogative constructions in signed languages: crosslinguistic perspectives. *Language* 80 (1), 7–39.

Cite this article: Cruz, M., & Frota, S. (2025). “Talking heads” in Portuguese sign and spoken languages, *Language and Cognition*, 17, e18, 1–21. <https://doi.org/10.1017/langcog.2024.63>