



Bauer, Anastasia & Gipper, Sonja & Herrmann, Tobias-Alexander & Hosemann, Jana. 2026. Rethinking linguistic feedback: A modality-agnostic and holistic approach to multimodal addressee signals in spoken and signed dyadic interaction. *Glossa: a journal of general linguistics* 11(1). pp. 1–50. DOI: <https://doi.org/10.16995/glossa.18539>

## Open Library of Humanities

# Rethinking linguistic feedback: A modality-agnostic and holistic approach to multimodal addressee signals in spoken and signed dyadic interaction

**Anastasia Bauer**, Department of Linguistics, University of Cologne, Germany, [anastasia.bauer@uni-koeln.de](mailto:anastasia.bauer@uni-koeln.de)

**Sonja Gipper**, Department of Linguistics, University of Cologne, Germany, [sonja.gipper@uni-koeln.de](mailto:sonja.gipper@uni-koeln.de)

**Tobias-Alexander Herrmann**, Slavic Institute, University of Cologne, Germany, [t.herrmann@uni-koeln.de](mailto:t.herrmann@uni-koeln.de)

**Jana Hosemann**, Sign Language Interpreting, University of Cologne, Germany, [jhoseman@uni-koeln.de](mailto:jhoseman@uni-koeln.de)

In this paper, we investigate multimodal recipient feedback in casual dyadic conversation in four languages: German Sign Language, Russian Sign Language, spoken German, and spoken Russian. Taking a modality-agnostic and holistic approach, we investigate the composition of conversational feedback from different multimodal signals, comparing sign and spoken languages without prioritizing any of the articulators or modalities. We find that in sign and spoken languages alike, feedback events include non-manual signals such as head movements or facial expressions in 85% or more of the instances. Across modalities, all four languages show only small percentages of feedback events without any non-manual elements, and head nods constitute the most frequent feedback signal. Moreover, we model three empirically observed *feedback styles* ranging from a style employing a rich array of non-manual signals, over one comprising mostly head movements, to a style relying somewhat more on the talk-oriented forms. Our data demonstrate that the basic infrastructure for feedback is shared among signers and speakers, while at the same time, signers and speakers show different probabilities for using one style or another. On the basis of these patterns, we propose a gradient model of feedback styles that generates testable predictions for future work. Our study emphasizes the importance of investigating interactional phenomena from a holistic, multi- and cross-modal perspective. As vocal and manual signals account only for a relatively small percentage of the feedback signals employed by the signers and speakers in our study, a linguistic theory that focuses solely on vocal and/or manual behavior remains incomplete and fails to account for the largest part of feedback in conversation. This study highlights that non-manual signals are fundamental to feedback and conversation more broadly,

*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by the Open Library of Humanities. © 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 OPEN ACCESS



and argues that theories of language must be reconceptualized, as purely speech-based accounts fail to capture the full complexity of human interaction. The findings have broader implications for theories of interaction and the Language Faculty: they underscore the need for models that integrate visual, non-manual, and interactional dimensions as constitutive elements of linguistic behavior. By highlighting the centrality of multimodal articulators in feedback production, this work contributes to a more comprehensive theory of human communicative interaction.

---

## 1 Introduction

When conversing, human interactants take turns (Sacks et al. 1974), thereby constantly changing roles between signer/speaker and recipient. While the current signer/speaker is the one who has the floor to convey some message to the recipient, the latter is by no means inactive. Rather, recipients in signed and spoken conversations are known to constantly provide the signer/speaker with feedback, for instance in the form of manual signs (e.g., YES), vocalizations (e.g., *mhm*), head nods, or smiles (Yngve 1970; Brunner 1979; Allwood et al. 1992; Bavelas et al. 2000; Gardner 2001; Cerrato & Skhiri 2003; Mesch 2016; Zellers 2021; Dingemanse et al. 2022; Lutzenberger et al. 2024). This is illustrated in (1)<sup>1</sup> taken from the DGS Corpus (Hanke et al. 2020) where signer B is nodding her head multiple times by way of providing feedback to A:

- (1) *DGS example illustrating an addressee head nod (DGS Corpus; Hanke et al. 2020). A short contextualized clip is available at <https://doi.org/10.6084/m9.figshare.30738701>, and the full video can be viewed in the DGS Corpus [here](#) (timestamp: 00:03:13.622).*

A: ICH ZIEL STEUER FACH STELLEN-VON-NACH BERATUNG  
 I AIM TAX SHELF TO-PUT-FROM-TO ADVICE  
 'It was my goal to become an assistant tax consultant  
 in the area of customer advisory service.'  
 B: multiple small head nods

---

<sup>1</sup> We acknowledge the challenges of providing traditional conversation analysis (CA) transcripts for the sign language examples in our study and commend recent research that has adopted more innovative and visually accessible methods, such as graphic comic-style representations (Skedsmo 2020; 2023). However, in our case, still images proved ineffective. The feedback signals we examine—such as slow, repeated head nods, subtle smiles, eyebrow raises or backward head movements—are often too subtle to be clearly conveyed in static images. Therefore, we have made our examples available online where possible and provide links to short video clips that illustrate the relevant feedback signals discussed in this paper. The videos of spoken German were collected in 2009 without participant consent for publication, and therefore cannot be made publicly available. Additionally, we provide multilinear notations inspired by sign language glossing conventions and using some conventions from the conversational transcription system GAT2 (Selting et al. 2009). In each example of dyadic interaction, the two signers/speakers are labeled as A and B. Following sign language annotation conventions, manual signs are glossed using small caps. If a single manual sign translates into multiple English words, these words are connected by hyphens. Non-manual signals are glossed with overlines, starting where the non-manual begins and ending where it ends. For DGS (German Sign Language), the glosses (as well as their translations) are based on the original corpus transcripts.

Feedback signals<sup>2</sup> fulfill a broad range of conversational functions: they may display the addressee's active participation, their understanding of the preceding contribution, or their capacity and willingness to go on with the conversation (Allwood et al. 1992). They may also give an evaluation of the content offered by the other interactant (Uhmann 1996), show affiliation with the signer/speaker (Stivers 2008), convey the presence or absence of conversational trouble (Schegloff 1982), and indicate whether a longer conversational unit is considered ongoing or completed (Koole & Gosen 2024). By providing feedback, recipients actively participate in shaping the signer's/speaker's talk (Tolins & Fox Tree 2014), where inadequate feedback can lead to the deterioration of the teller's performance (Bavelas et al. 2000) or to the production of repair sequences (Byun et al. 2018). Moreover, the way a person gives feedback influences how others perceive their personality (Blomsma et al. 2022), and there is even some evidence that the feedback style of a person indeed shows a relationship to their personality traits (Bendel Larcher 2021).

All this suggests that feedback constitutes a vital mechanism in human communication and cognition. Given its centrality, feedback can provide us with a window into the mechanisms of human social interaction in general. Any theory of language and communication must therefore be able to account for feedback phenomena in conversation.

The forms of feedback signals, their frequencies as well as their typical conversational employment vary across languages, across varieties of the same language, and even across individuals using the same language (White 1989; Maynard 1990; Tottie 1991; Clancy et al. 1996; Stubbe 1998; Dideriksen et al. 2023; Blomsma et al. 2024). However, striking similarities have also been noted, both among spoken languages as well as between sign and spoken languages. In spoken languages, the employment of a form containing a nasal consonant such as *mhm* is pervasive across languages from different families and with different typological profiles (Dingemans et al. 2022). Regarding similarities between sign and spoken languages, feedback signals in both British Sign Language and British English have been shown to rely to a great extent on head movements (Lutzenberger et al. 2024).

With language existing in at least three combinations of modalities—spoken/auditory, signed/visual, and signed/tactile—there is a strong need to study feedback across modalities using similar approaches. Today it is widely acknowledged that human language in its primary co-present context is a fundamentally multimodal phenomenon (Goodwin 1986; Bavelas 1990; Kendon 2004; Vigliocco et al. 2014; Abner et al. 2015; Mondada 2016; Keevallik 2018; Perniss 2018; Holler & Levinson 2019; Özyürek 2021; Rasenberg et al. 2022; Kendrick et al. 2023; Sandler

---

<sup>2</sup> There are certain elements of feedback that may be intentionally produced in some instances and unintentionally produced in other instances (e.g., smiles, eyebrow flashes). Here we consider all instances of such elements to be signals, and do not attempt to differentiate intentional from unintentional signals.

2024) that involves the coordination of various articulators (Holler & Levinson 2019; Özyürek 2021). The same is true for feedback (Cassell & Thorisson 1999; Allwood & Cerrato 2003; Allwood et al. 2007a; b; Bertrand et al. 2007; Navarretta & Paggio 2010; Truong et al. 2011; Navarretta & Paggio 2012; Malisz et al. 2016; Rasenberg et al. 2022; Boudin et al. 2024). Nevertheless, research adopting a holistic multimodal perspective on feedback remains scarce, and our understanding of how vocal and visual, manual and non-manual signals combine into complex recipient feedback in everyday conversation is still limited, particularly comparing sign and spoken languages.

With the current study, we address this gap by investigating how feedback varies in form and frequency in signed and spoken languages, expanding upon previous observations in the literature. We examine feedback from a multimodal and cross-linguistic perspective by utilizing corpora of casual conversations from four different languages: German Sign Language (DGS), Russian Sign Language (RSL), spoken German (GER), and spoken Russian (RUS) (Hoffmann & Himmelmann 2009; Burkova 2015; Konrad et al. 2020; Bauer & Poryadin 2023; Bauer 2023). Our focus is on possible compositions of what we call *feedback events*, which consist of multiple signals produced with different articulators (see Section 3.2). We take into account various signals which can be produced during feedback, including, e.g. words like *ja* ‘yes’, manual signs such as DGS STIMMT ‘right’ or mouthings like *okay*<sup>3</sup>, vocalizations like *mhm* and non-manual signals such as head nods, eyebrow raises, smiles and others. Using parallel annotation and analysis, we annotated at least 43 minutes of co-present dyadic conversations<sup>4</sup> in each of the four languages and identified ca. 1,900 instances of feedback. Crucially, our approach is modality-agnostic, meaning that we analyze feedback signals and events without privileging one articulator over another. We are inspired by the work of Hodge et al. (2023), who conducted a modality-agnostic comparison of quotatives in Auslan (Australian Sign Language) and the spoken language Matukar Panau (Oceanic). While Hodge et al. (2023) examined different articulators available in sign and spoken languages (e.g., mouthing only in Auslan, speech only in Matukar Panau), the present study extends this approach by introducing a methodological framework that enables an integrated comparison of these types of signals. Specifically, we group two articulators together (hands and mouth) and classify manual signs, mouthings, spoken words, and vocalizations under the unified category of *talk*. This allows us to compare sign and spoken languages without overemphasizing differences imposed by the constraints of their respective modalities—a limitation also identified by Hodge et al. (2023)<sup>5</sup>.

---

<sup>3</sup> Mouthings are mouth movements produced during sign language interaction which resemble words (or parts of them) from the surrounding spoken/written language (Bauer & Kyuseva 2022).

<sup>4</sup> We follow Ameka & Terkourafi (2019) and use the term “co-present dyadic conversation” to refer to what is often called “face-to-face” interaction. We use it to describe interactions between two people who share the same physical space, while recognizing that cultural norms, such as avoiding eye contact, can influence the shape of social interaction (Rossano et al. 2009).

<sup>5</sup> For more information on the contents of the category of *talk*, see Section 3.2.

Our data show similarities between signed and spoken language modalities in the architecture of feedback events, as most feedback events (85% or more) involve non-manual signals such as head and/or facial movements in all four languages. These feedback events were produced either non-manually alone or in combination with signed/spoken elements across languages. Moreover, in all four languages, the most frequent feedback event design is that of a multiple head nod without any additional signal, emphasizing the importance of head nods in feedback in the four languages examined. We interpret these findings as contributing to the accumulating evidence supporting the existence of a shared interactional infrastructure of conversation among both signers and speakers (Lutzenberger et al. 2024).

Despite frequent reference to multimodality in contemporary linguistic discourse, the full complexity of multimodal human communication remains largely underrepresented in many prevailing linguistic and cognitive theories, which often rely on unimodal conceptions of language. However, there are notable exceptions, such as the recent work by Cohn & Schilperoord (2024). A comprehensive linguistic theory must account for language as a multimodal system, encompassing both the vocal-auditory and gestural-visual articulators, and must situate these within the broader framework of human cognition.

## 2 Previous research on multimodal feedback in sign and spoken languages

Conversational feedback has been referred to by various terms in the literature, the most common being *backchannels* (Yngve 1970), *listener* or *minimal responses* (Hess & Johnston 1988; Bavelas et al. 2002; Fujimoto 2009), and *reactive* or *response tokens* (Gardner 2001; McCarthy 2003; Xu 2016). The term *feedback*, as employed in this study, was introduced by Allwood et al. (1992). Despite ongoing terminological differences (Simon 2018), there is a general agreement among researchers that feedback signals must be distinguished based on the pragmatic or communicative functions they serve. For instance, some utterances may signal active participation, others may acknowledge and agree with what has been stated, while others might treat new information as newsworthy or provide an evaluative comment (see **Figure 1**).

Prior to 2000, research predominantly focused on vocal responses, such as *mm*, *yeah*, or *okay*, primarily in spoken English (Beach 1993; Drummond & Hopper 1993; Jefferson 1993). However, some researchers recognized that feedback encompassed more than oral behavior, drawing attention to visual signals. Dittmann & Llewellyn (1968) are among the first to acknowledge the relationship between vocal responses and head nods during feedback, and Yngve (1970) already emphasizes the importance of investigating video instead of audio data for the study of feedback in conversation. Brunner (1979) and Jefferson (1984) highlight smiles and laughter in conversation, while other linguists include various head movements in interactions in the same category as vocal expressions like *uh*, *yeah*, and co-completions or requests for clarification (Kendon 1967;

Duncan 1974; Hadar et al. 1985). These studies initiated a tradition of studying feedback from a multimodal perspective.

Although the potential of (combining) visual signals in feedback-giving is vast, most research has concentrated on individual feedback signals from a single articulator (Allwood & Cerrato 2003; Bertrand et al. 2007; Hömke et al. 2017; Kendrick & Holler 2017). While some studies recognize the role of various articulators, such as for example head movements and smiles, they often do not integrate these into a holistic analysis of feedback (Bavelas et al. 2000; Gardner 2001; Lindblad & Allwood 2015; Gironzetti et al. 2016; Malisz et al. 2016). Few studies have concurrently examined multiple feedback signals. Blomsma et al. (2024), for example, analyze a variety of facial gestures across multiple addressees, but their study is limited to a single spoken language and does not involve real human–human interaction.

In comparison to research on spoken languages, studies on feedback mechanisms in sign languages remain relatively scarce. Existing literature has primarily focused on repair mechanisms, documenting them in Argentine Sign Language (Manrique & Enfield 2015; Manrique 2016), Swiss German Sign Language (Girard-Groeber 2015), Norwegian Sign Language (NTS) (Skedsmo 2020), Balinese homesign (Safar & De Vos 2022), Providence Island Sign Language (Omardeen 2023), British Sign Language (BSL) (Lutzenberger et al. 2024), and in cross-signing contexts, where Deaf signers of different sign languages meet for the first time (Byun et al. 2018).

With respect to non-repair feedback, there is much less research on sign languages. Backer (1977) offers a brief description of what she terms *regulators* in a small corpus of American Sign Language (ASL), building upon the work by Wiener & Devoe (1974) who made a systematic description of those behaviors in the visual, vocal, postural, and gestural articulators that signal and/or monitor the initiation, maintenance, and termination of spoken messages. Backer (1977) differentiates between feedback signals that initiate a turn (such as an increase in size and quantity of head-nodding, movement of hands out of rest position, i.e. indexing, touching, or waving hand in front of interlocutor, gaze) and feedback signals produced in passive reciprocity (gaze, head nodding, smiling, postural shifts, facial activity expressing surprise, agreement, uncertainty, lack of understanding, etc.) or short repetitions of some of the interlocutor’s signs. Subsequent research by Coates & Sutton-Spence (2001) further classified turn-taking regulation in sign language interactions into two categories: non-manual and manual. This distinction seems important for future research, as non-manual elements rather than manual signs appear to play a more critical role in conveying feedback (Lutzenberger et al. 2024), a pattern also observed in the present study.

Mesch (2016) reports for the first time on backchannels signals in Swedish Sign Language (STS), noting that manual backchannels (such as palm-up, YES, INDEX, AGREE, EXACTLY) are quite rare and often produced in the signer’s lap. STS signers predominantly use

non-manual backchannel signals such as nodding, head-shaking, smiling, changes in posture, nose wrinkling, or widened eyes to signal feedback. In her analysis of 35 minutes of STS dialogues involving 16 Deaf signers, Mesch (2016) focuses primarily on manual backchannels, which generally consist of one to three signs/gestures in STS, with palm-up and YES being particularly frequent.

A recent study by Börstell (2024) also focuses on manual feedback in STS, specifically on continuers. Using the approach proposed by Dingemanse et al. (2022), Börstell (2024) examines continuer candidates within a subset of the STS corpus. This study supports the findings by Mesch (2016) that the two manual elements YES and palm-up are the most frequent manual backchannels in STS. Similar to Mesch (2016), Börstell (2024) excludes non-manual signals due to the limited annotation of non-manual expressions in the dataset.

Fenlon et al. (2013) examined gender and age differences in turn length and the frequency of backchannels in BSL dyadic conversations. Contrary to earlier studies on spoken languages (Duncan 1974; Bilous & Krauss 1988), they found no significant differences between gender and age groups in the time spent on manual or non-manual backchannels.

Lutzenberger et al. (2024) provide the first cross-linguistic comparison of feedback in a signed and a spoken language. Their recent study on repairs and continuers in BSL and British English revealed similarities in discourse management strategies among signers and speakers who share a common cultural background. They observe that the interactional infrastructure used by both signers and speakers predominantly relies on behaviors of the head, face, and body—alone or combined with what they call ‘verbal’ elements (spoken words or manual signs)—while solely ‘verbal’ strategies are rare, similar to what was found by Mesch (2016) earlier.

In DGS (German Sign Language), head nods have been found to play a crucial role in interaction and even exhibit distinct phonetic characteristics when used as feedback. This was demonstrated in a recent study by Bauer et al. (2024), who used OpenPose to analyze the kinematic properties of head nods, revealing that feedback nods are slower and smaller in amplitude than affirmative nods.

However, research on feedback in signed conversations is still in its early stages. This may be due, in part, to a longstanding manual bias in the study of sign languages (Puupponen 2019). Sign language linguistics has largely focused on lexical, phonological, and morpho-syntactic structures, often overlooking the interactive dimensions of communication (Lepout & Shaw 2022). Yet interaction consists of composite utterances (Kendon 2004), in which non-manual actions combine with manual and/or vocal actions—a perspective that has received little attention in sign language research to date.

The majority of all existing studies on feedback in sign languages has concentrated on manual backchannels, also due to the challenges of annotating non-manual signals. Additionally, most

research has focused on continuers, as these are more readily identifiable compared to other feedback types. Our study seeks to address this gap by examining the full range of multimodal feedback, encompassing various feedback types produced by different articulators (see Section 3.1 for an explanation of the various types of feedback).

### 3 The current study

The literature summarized in Section 2 suggests that there are striking similarities between signed and spoken languages with respect to the composition of feedback, in that head movements are very frequently involved in both types of languages. However, differences are also apparent: where spoken languages employ speech (including both lexical and non-lexical tokens such as *yeah* or *mhm*), feedback in sign languages sometimes contains signs such as YES or gestures as palm-up which are often signed at a location low in the signing space (Mesch 2016; Börstell 2024). The existence of nasal feedback signals such as *mhm* in spoken languages and low-signed signals in sign languages suggests that speakers and signers alike strive to minimize the effort of production and reduce the potential intrusiveness of feedback (Dingemanse et al. 2022; Lutzenberger et al. 2024; Börstell 2024).

However, the differences summarized above are in fact due to constraints of the particular modality (see also Vandenitte 2023). Although sign language users may potentially employ nasal vocalizations, these are not visually perceptible. Likewise, it is obvious that spoken languages do not possess lexical manual signs. How, then, can we compare sign and spoken languages in a meaningful way without overemphasizing these modality-based constraints? We suggest that this is possible with a modality-agnostic (Hodge et al. 2023) approach to feedback (see Section 3.2).

Also, in previous research on conversational feedback, the focus often lies on one single type of feedback signal, e.g., manual signs and gestures (Mesch 2016), vocalizations (Gardner 2001; Zellers 2021), nodding (Stivers 2008), or smiles (Brunner 1979). More holistic approaches looking at the multimodal composition of feedback are rarer (Lindblad & Allwood 2015; Lutzenberger et al. 2024). Our study is grounded in such holistic approaches, as our multimodal and modality-agnostic perspective on feedback assumes that feedback can consist of multiple layers conveyed through different articulators.

In order to develop such an approach, in Section 3.2 we redefine feedback in a way that allows us to investigate it from two perspectives: a holistic perspective, taking the composition of instances of feedback into account, and an atomic perspective looking at different articulators involved in formulating feedback separately. Before we start presenting our modality-agnostic and holistic approach, in Section 3.1 we briefly discuss the delimitation of the phenomenon of feedback as implemented in our study.

### 3.1 Delimiting feedback

In order to make transparent which kinds of signals we include in our study, in the following we propose a schema of conversational feedback that allows for an integration of different types of signals found in different languages, while at the same time making clear distinctions between signals that are oriented toward indicating the presence vs. the absence of trouble.

The first important distinction we draw is between feedback and other kinds of responses in interaction. Feedback events are fundamentally responsive in that they are by definition related to some preceding talk by another interactant. However, it is crucial to distinguish feedback from other types of responses: while feedback events can be and are regularly solicited by the current signer or speaker, it is crucial that they are not made conditionally relevant in the same way as, for instance, answers to questions. This means that, in principle, the recipient can decide about the placement of feedback signals and may even withhold relevant feedback (Schegloff 1982: 86). Although the lack of feedback can lead to conversational failure (Bavelas et al. 2000), which is also the case for conditionally relevant responses in that they are, if withheld, “officially absent” (Schegloff 1968: 1083), the consequences are different: Conditionally relevant responses—answers to questions, acceptance of invitations, etc.—are responses that are specifically relevant at the particular point in time, with restrictions on possible responses set by the initial utterance. Feedback, in contrast, may be relevant at any point in a conversation, due to the fact that communicative trouble may arise at any time and, consequently, make the initiation of repair—or the passing of that opportunity—necessary (Schegloff 1982). Thus, the conditions on sequential placement are very different for conditionally relevant responses and feedback. This is illustrated by Example (2) showing a question–answer pair, and Example (3) showing the use of a feedback signal. In (2), interactant A poses a polar question in the first line, which interactant B answers in the second line. The question restricts B’s possibilities for responding, as an answer of the type yes/no is made conditionally relevant by the question. B responds to the question by means of the response particle *ja* ‘yes’.

- (2) *Spoken German example illustrating a question–answer sequence from MünsterKorpus\_DB (Hoffmann & Himmelmann 2009).*

A: hat er dir unterschrie[ben?  
 has he you signed  
 ‘He signed [the poster] for you?’

B: [ja::;  
 yes  
 ‘Yes!’

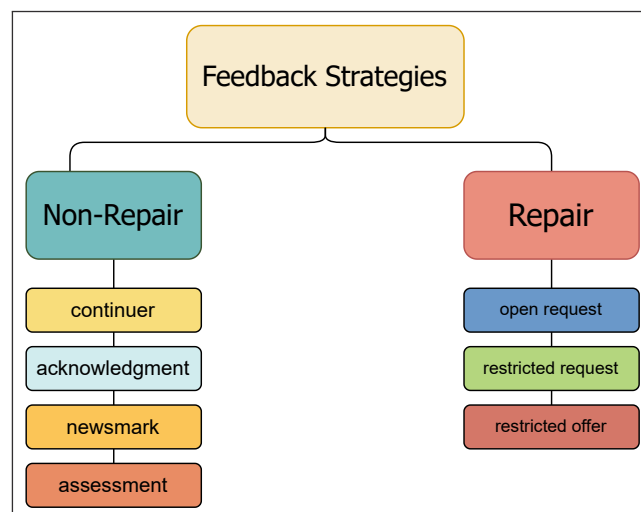
In (3), in contrast, A offers a piece of information to B in the first line. This action does not make any particular response relevant, so B can choose to provide feedback or withhold it. The possibilities for responding are thus not restricted in the same way as by the question in (2). This does not mean that feedback may not be preferred over silence in this context, it just means that it is not conditionally relevant in the same way as an answer to a question. Speaker B can also choose the type of feedback she provides. In this case, she chooses a verbal feedback token in the form of the response particle *ja* 'yes'. Here we can observe the multifunctionality of response particles: *ja* 'yes' is employed to formulate an affirmative answer to a polar question in (2), and as a continuer in (3).

(3) *Spoken German example illustrating a sequence with feedback from MünsterKorpus\_DB (Hoffmann & Himmelmann 2009).*

A: un dann ham wir einfach äh n teppich gekauft.  
 and then have we simply uhm a carpet bought  
 'And then we just bought a carpet.'

B: (0.6) **ja** ;  
 yes  
 'Yes.'

Another important distinction to draw is that between different types of feedback, based on whether it deals with some conversational trouble (repair) or indicates a lack thereof (non-repair feedback) (see **Figure 1**).



**Figure 1:** Feedback strategies.

**Figure 1** illustrates various feedback strategies. Although repair mechanisms (Dingemanse & Enfield 2015; Dingemanse et al. 2015) are included in the figure for the sake of completeness, they are excluded from further discussion, as they are not the focus of this study. Instead, we investigate conversational moves that do not initiate or constitute repair, but rather imply that repair is unnecessary. One of the most well-known and widely studied types of feedback is what is often referred to as a continuer. Continuers convey at least the basic interactional function of passing on the opportunity for initiating repair (Schegloff 1982)—see Example (3) above.<sup>6</sup> Moreover, feedback signals called ‘newsmarks’, in addition, explicitly treat the information given by the preceding interactant as new and mark it as ‘remarkable’ (Marmorstein & Szczepek Reed 2023). In this category, we include non-repetitional requests for reconfirmation (Gipper et al. 2023) such as German *echt?* ‘really?’, see Example (4), as well as change-of-state tokens (Heritage 1984) such as *ah*, see Example (5).

(4) *Spoken German example of a newsmark from MünsterKorpus\_DB (Hoffmann & Himmelmann 2009).*

A: mein papa is gro:ßer HSV-f[an-  
 my dad is big fan\_of\_HSV  
 ‘My father is an enthusiastic fan of HSV.’

B: wide open smile  
 eyebrow raise  
 eyes wide open  
 head forward  
**[echt?**  
 really  
 ‘Really?’

A: wide open smile  
multiple small nods  
 (.) ja-  
 yes  
 ‘Yes.’

---

<sup>6</sup> Some scholars propose a distinction between continuers and acknowledgments, where acknowledgments indicate agreement with or understanding of the previous turn (Gardner 2001: 2). In this study, we do not distinguish between continuers and acknowledgments, as they often share similar forms. Moreover, it is notoriously difficult to determine whether an interactant aims to indicate agreement or just pass on the opportunity for repair. We suggest that a future operationalization of the distinction between the two could be conceptualized on the basis of their sequential position: continuers typically follow a volunteered initial utterance (second position), while acknowledgments follow a conditionally relevant response (third position). We leave an evaluation of this proposal for future research.

- (5) *Spoken Russian example illustrating a newsmark from [Russian Multimodal Conversation Corpus](https://doi.org/10.6084/m9.figshare.30738701) (Bauer & Poryadin 2023). A short clip is available at <https://doi.org/10.6084/m9.figshare.30738701>.*

A: da èto stipendija Bundestaga vot=  
 yes this scholarship Bundestag PARTICLE  
 'Yes, this is a Bundestag scholarship.'

B: head back  
 =ah: :

Lastly, there are feedback signals that overtly indicate some kind of evaluation of the preceding information, 'assessments' (Uhmman 1996), as in Example (6).

- (6) *RSL example illustrating a feedback event comprising a manual sign and multiple small head nods functioning as an assessment ([the RSL Corpus](https://doi.org/10.6084/m9.figshare.30738701); Burkova 2015). A short contextualized clip is available at <https://doi.org/10.6084/m9.figshare.30738701>, and the full video can be viewed in [the RSL Corpus](https://doi.org/10.6084/m9.figshare.30738701) after registration (RSLN-d-s23-s24, timestamp 00:00:05.965).*

A: INDEX ODNOKLASSNIK ITOG DEVJAT' ČELOVEK [DEVJAT'] ČELOVEK=  
 INDEX CLASSMATE TOTAL NINE PERSON NINE PERSON  
 'I had nine classmates in total, nine persons.'

B: multiple head nods  
 [DEVJAT' ČELOVEK]  
 'Nine persons.'

A: =[DAVNO POMNIT' DA  
 LONG-AGO REMEMBER YES  
 'It is long ago, but I still remember it.'

B: multiple small head nods  
 [INTERESNO  
 'Interesting.'

We include these three types of feedback events—continuers, newsmarks, and assessments—in our investigation, regardless of their sequential position (second, i.e., following a volunteered initial utterance, or third, i.e., following a response made conditionally relevant) or their turn-taking properties (passive reciprocity vs. incipient speakership, see Sbranna et al. 2022).

For this study, we did not include repetitions (see, e.g., the first part in Example (6)), as at least for German it has been shown that they tend to fulfill relatively marked actions when used in requests for reconfirmation (Gipper et al. 2023). Given that it is not clear whether this is also true for the other three languages, we chose to exclude repetitions and leave their investigation for future research, as they may not be fully comparable across the languages in our sample.

### 3.2 A modality-agnostic and holistic approach to feedback

In this paper, we take a modality-agnostic (Hodge et al. 2023) approach to the comparison of sign and spoken languages, an approach that looks at all components of a feedback event without privileging any of the articulators. Most of the signals functioning as feedback in our study, produced by the various articulators—head, eyebrows, eyes, nose, mouth gestures, cheeks, manual gestures, and shoulders—are comparable across sign and spoken languages. We thus observe that their use in feedback in signed languages is not qualitatively different from that in spoken languages. While we acknowledge that the meanings in feedback need not be the same for all articulators in the four languages, we start out with a descriptive and exploratory approach allowing for the possibility that signals are used in similar ways across languages. A detailed analysis of the exact meanings in feedback and other interactional functions will be an intriguing topic for future research. Our data show, for example, that both eyebrow raises and nose wrinkles are used in all four languages. For eyebrow raises, we can say that they are used in very similar ways in a newsmarking function, albeit with different frequencies. With regard to nose wrinkling, DGS shows markedly more nose wrinkles than the other languages (8% vs. 0.5% or less) (see **Table 8**). A nose wrinkle is known to convey the meaning ‘that’s right’ in DGS interaction (Herrmann 2020), but its use in spoken German discourse has not been addressed in the literature. A fuller analysis of this difference is left for future research. Moreover, the plurifunctionality of non-manual gestures is a well-established phenomenon (Andries et al. 2023; Oomen & Roelofsen 2023), and we acknowledge that many of the signals examined in this study may serve multiple functions simultaneously.

At the same time, we recognize that some articulators may be argued to differ across sign and spoken languages in the production of feedback signals. In sign languages, feedback may be conveyed through manual signs (see DGS Example (7) or RSL Example (6)) as well as mouthings (see DGS Example (7) and RSL Example (8)), mouth movements resembling spoken or written forms of the surrounding language (Bauer & Kyuseva 2022). In contrast, spoken languages express feedback through spoken words (see Example (3)) or vocalizations (e.g., *mhm*) (Dingemanse et al. 2022).

- (7) *DGS example illustrating a feedback event containing a head nod, a manual sign and a mouthing (DGS Corpus; Hanke et al. 2020). A short clip is available at <https://doi.org/10.6084/m9.figshare.30738701>, and the full video can be viewed in the DGS Corpus, timestamp 00:10:39.304).*

A: LIST AUSTRALIEN LIST SÜDEN AMERIKA ARGENTIEN  
 LIST AUSTRALIA LIST SOUTH AMERICA ARGENTINA  
 'I've been to Australia, South America, Argentina.'

B: head nod multiple head nods  
/ah/ /schön/  
**SCHÖN**  
 NICE  
 'Ah, nice.'

- (8) *RSL example illustrating a feedback event containing a head nod and a mouthing* (from the [RSL Conversations Corpus](https://doi.org/10.6084/m9.figshare.30738701); Bauer & Poryadin 2023). A short clip is available at <https://doi.org/10.6084/m9.figshare.30738701>.

A: POTOM PRIECHAT' AVSTRIJA INDEX POTOM=  
 THEN COME AUSTRIA INDEX THEN  
 'Then to go to Austria.'

B: small nod multiple nods large nod  
/da da/  
 'Yes, yes.'

A: =POECHAT' BERLIN B-E-R-L-I-N INDEX  
 GO BERLIN B-E-R-L-I-N INDEX  
 'Then to Berlin.'

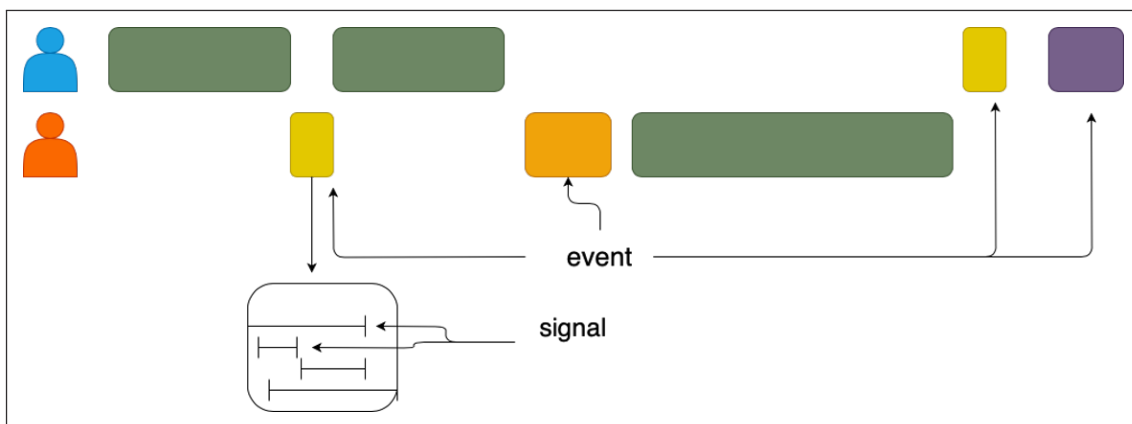
B: large nod

Classifying these signals on the basis of the articulators involved (hand, mouth, and mouth, respectively), would obscure the fact that these differences are based on modality-specific constraints for the two types of languages. Therefore, rather than classifying these three types of signals on the basis of the articulators with which they are produced, we classify them as one single category which we call *talk*<sup>7</sup> for all four languages, signed and spoken. This allows us to compare sign and spoken languages with respect to the extent to which they employ *talk* elements regardless of the modality.

---

<sup>7</sup> Lutzenberger et al. (2024) employ the term 'verbal' to include manual BSL signs and spoken English words. We use *talk* as an alternative to 'verbal' in this context, as 'verbal' typically carries connotations specific to spoken language. Moreover, we include mouthings, as these clearly convey lexical (e.g., *ja* 'yes') and non-lexical content (e.g., *ah*).

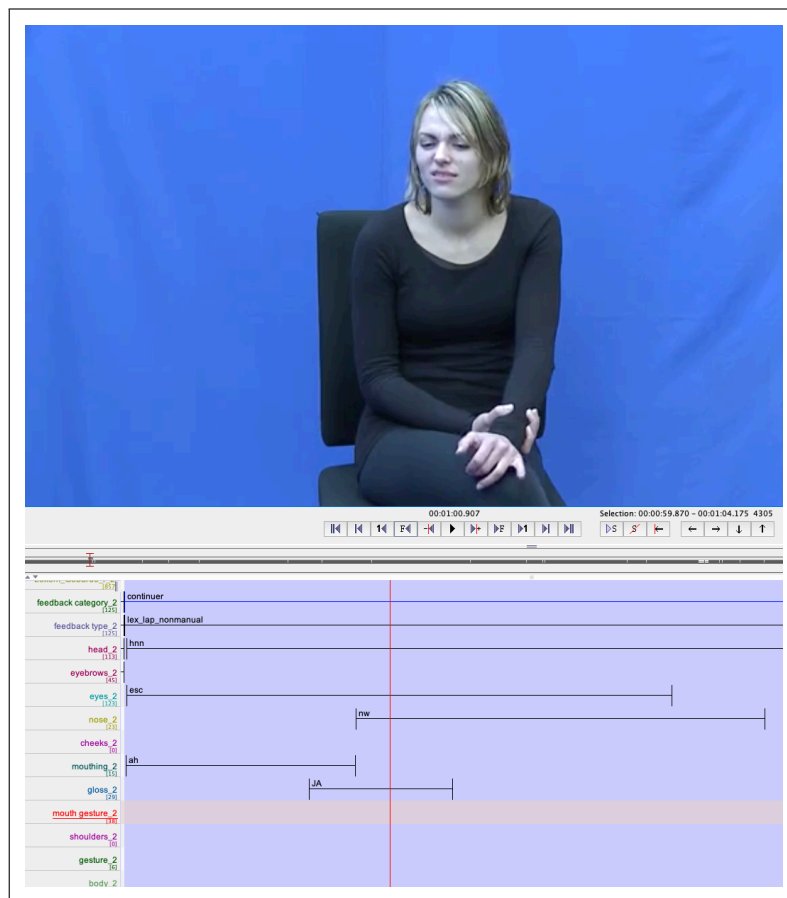
In addition to this modality-agnostic approach, in the following we propose a novel *holistic* construal of conversational feedback, where different articulators are employed to produce signals that combine into what we call a *feedback event*—see **Figure 2**. The figure shows a stretch of conversation between two interactants. The dark green longer bars represent turns, while the shorter bars indicate feedback events. Different colors represent different types of feedback, such as continuers or assessments, with varying durations. The zoom-in window illustrates how various signals may be combined within a single feedback event—some lasting the full duration of the event, others occurring only briefly.



**Figure 2:** Feedback signals vs. feedback events.

As can be seen in the ELAN screenshot in **Figure 3**, the person produces distinct signals such as a manual sign JA ‘YES’ in her lap, a mouthing *ah*, a head nod (*hnn*), squinted eyes (*esc*) and a nose wrinkle (*nw*) with different articulators, such as the head, the eyes, the mouth and the nose. These collectively form a *feedback event*. While some feedback events may consist of a single signal, this study shows that they often comprise multiple simultaneous signals (see **Figure 7**). So, instead of analyzing one signal, such as head nods, we claim that it is essential to take all (potential) articulators into account to get a broader understanding of the composition of feedback events. Crucially, this perspective does not entail that all signals that constitute a feedback event necessarily convey one single meaning. Rather, our approach focuses on temporal aspects of co-occurrence.

This redefinition allows us to investigate feedback from two perspectives, looking at the whole, potentially multi-layered feedback event, but also at its components. Moreover, it allows for a modality-agnostic approach to feedback where the employment of different articulators is compared across sign and spoken languages.



**Figure 3:** A screenshot from ELAN showing an example of a multilayered feedback event in DGS (source: DGS corpus, Hanke et al. 2020, file koe\_03\_sachgebiete).

### 3.3 Research questions

In this study, we aim to contribute to our knowledge of the similarities and differences in the composition of feedback events between sign and spoken languages. For this purpose, we compare conversational feedback in four languages, two signed and two spoken, matched according to cultural background: German Sign Language (DGS), Russian Sign Language (RSL), spoken German (GER), and spoken Russian (RUS). We annotated feedback events in corpora of casual conversations according to our definition in Sections 3.1 and 3.2, developing a coding scheme based on previous research and our own findings during annotation (see the Appendix).

In order to investigate possible similarities and differences between the four languages, we employ a descriptive and exploratory approach. Research comparing sign and spoken languages is still too scarce to formulate meaningful hypotheses. Moreover, the formulation of hypotheses would pose unhelpful restrictions on our research, where an exploratory approach opens the possibility for unanticipated findings.

We start our research with the following questions:

- I. What are the typical components of feedback events across languages?

Previous research suggests that signed and spoken languages both rely to a large extent on head movements when expressing feedback events (Lutzenberger et al. 2024). Moreover, we know that non-manual signals such as head nods can combine with other signals in order to formulate multi-layered feedback events (Dittmann & Llewellyn 1968). Building on this research, we aim to further investigate the components involved in the formulation of feedback events, with special attention paid to similarities and differences between sign and spoken languages.

- II. What is the role of language, language modality, cultural background, and individual signer/speaker in the formulation of feedback events?

As initial research suggests that the formulation of feedback events may in fact be quite similar in sign and spoken languages (Lutzenberger et al. 2024), we use our matched datasets to look at the question of whether feedback producers of one language are more similar to each other than to those of another language, or whether similarities and differences can rather be explained in terms of cultural background or language modality.

## 4 Materials and methods

### 4.1 Data

For all four languages, we investigated video-recordings of free mundane dyadic conversations, using four (DGS, RSL, RUS, GER) corpora. For each language, we included three such conversations. In almost all dyads, both participants were familiar with each other prior to the recording.<sup>8</sup> In the Russian dataset, however, two participants met for the first time on the day of the recording. In the same dataset, one interactant features in two recordings, so we investigate data from five interactants for Russian and from six interactants for the other three languages. For each language, we annotated between 43 and 58 minutes of conversation. In line with our exploratory approach, we chose to annotate similar amounts of time in order to be able to compare languages with respect to how frequent feedback is.

Our DGS data has been taken from the Public DGS Corpus (Hanke et al. 2020). The DGS Corpus is an annotated reference corpus of German Sign Language, 50 hours of which have been made publicly available. Its 330 participants use DGS as their primary language of daily life and come from various regions of Germany (Schulder & Hanke 2022). The DGS content analyzed

---

<sup>8</sup> Information on whether interlocutors knew each other is not publicly available in the DGS Corpus metadata. We therefore selected interactions that, based on their content, strongly suggested familiarity between the participants—for example, references to mutual friends, shared holidays, or inquiries about each other's partners.

and presented in this paper is drawn from release 3 of *MY DGS – annotated* (Hanke et al. 2020; Konrad et al. 2020), a research dataset that provides Public DGS Corpus recordings with full sign annotations and translations in German and English.

Our RSL data come from two corpora. One file was sourced from the RSL Online Corpus, developed by Svetlana Burkova and her team at Novosibirsk University (Burkova 2015). This corpus currently comprises over 230 recordings from 43 RSL signers (both men and women, aged between 18 and 63) including Deaf and Hard-of-Hearing individuals. To obtain authentic conversational data, we selected a 20-minute unprompted conversation between two Deaf signers recorded in Novosibirsk.

Due to the limited amount of interactional data in this corpus, we also used a second corpus of RSL conversations. The additional two casual dyadic conversations, lasting between 40 and 60 minutes, were sourced from Bauer & Poryadin (2023) and feature Deaf native RSL signers. This recently compiled corpus includes data from signers who previously lived in St. Petersburg, Čita, and Černišov before immigrating to Germany. Participants discussed a range of topics, including their lives in Russia before immigrating and their experiences as Deaf individuals in Europe and Russia.

Our German data constitute a subset of the Münster Korpus (Hoffmann & Himmelmann 2009), an unpublished corpus of video-recorded conversations among students from different areas in Germany who speak Standard colloquial German as their first language. The conversations were recorded in the year 2009 in the city of Münster, Germany.

Our three spoken Russian conversations are part of a Russian Multimodal Conversation Corpus (Bauer & Poryadin 2023). This corpus features dialogues among Russian immigrants lasting 40–60 minutes each. Participants, aged 20–30 years old, are native Russian speakers residing in Germany for no longer than 5 years.<sup>9</sup>

We aimed to use data that were maximally comparable across both interactional type (free, unprompted conversations between acquainted interlocutors) and recording setup (two participants). **Table 1** summarizes the sources, the amount of time and tokens for the data employed in this study. **Table 2** lists details for the different recordings.

---

<sup>9</sup> We acknowledge that the Russian as well as RSL data in this study stems primarily from Russian interactants living in Germany, which may raise questions about cultural generalizability. This is a valid concern and one we carefully considered during data collection. To address this, we selected participants who use RSL/Russian in their daily lives, particularly in interactions with family and friends. For example, one signer reported communicating exclusively with RSL-signing friends. Additionally, both the RSL and spoken Russian data were collected from individuals who migrated to Germany within the past two to five years. As no multimodal corpus of dyadic interaction currently exists for spoken Russian, and the available RSL corpus from Russia (Burkova 2015) includes only a limited amount of free interaction, we opted to collect new data in Germany. Fieldwork in Russia was not feasible after 2022. Despite these circumstances, we believe the data remain valid and representative for our research purposes.

LANG.	SOURCE	INTERACTANTS	MIN.	EVENTS
DGS	Hanke et al. (2020); Konrad et al. (2020)	3 f, 3 m	48	585
RSL	Burkova (2015); Bauer & Poryadin (2023)	2 f, 4 m	43	397
GER	Hoffmann & Himmelmann (2009)	3 f, 3 m	45	525
RUS	Bauer & Poryadin (2023)	4 f, 1 m	58	419

**Table 1:** Summary of the data employed in this study, including language, sources, durations of the annotated data, number and gender of interactants, and counts of feedback events.

LANG.	TRANSCRIPT	AGE	GENDER	MIN.	SOURCE
DGS	koe_01_free_conversation	18–30	fm	14:51	<a href="https://doi.org/m59x">https://doi.org/m59x</a>
	koe_03_sachgebiete	18–30	fm	15:25	<a href="https://doi.org/m59z">https://doi.org/m59z</a>
	koe_04_free_conversation	18–30	fm	18:10	<a href="https://doi.org/kz87">https://doi.org/kz87</a>
RSL	RSLC_s3_s4_180423	31–45	fm	09:59	<a href="https://doi.org/npzp">https://doi.org/npzp</a>
	RSLN_d2_s8_s9	31–45	fm	19:52	<a href="http://rsl.nstu.ru">http://rsl.nstu.ru</a>
	RSLC_s1_s2_180423	60+	mm	14:43	<a href="https://doi.org/npzp">https://doi.org/npzp</a>
GER	MünsterKorpus_UV	18–30	fm	16:30	
	MünsterKorpus_DB	18–30	fm	16:22	
	MünsterKorpus_LD	18–30	fm	13:07	
RUS	RCC_s1_s11_010923	18–30	fm	27:00	<a href="https://doi.org/npzr">https://doi.org/npzr</a>
	RCC_s12_s10_010923	18–30	ff	14:49	<a href="https://doi.org/npzr">https://doi.org/npzr</a>
	RCC_s1_s2_010923	18–30	ff	16:00	<a href="https://doi.org/npzr">https://doi.org/npzr</a>

**Table 2:** Overview of annotated transcripts.

## 4.2 Annotations

All data were annotated in ELAN (The Language Archive, MPI Nijmegen, The Netherlands; e.g. Crasborn & Sloetjes 2008). We started out with an annotation scheme developed by the first author inspired by the RSL Corpus annotations (Burkova 2015). For each feedback event, we created an annotation on a separate tier in ELAN marking the length of the whole event. A feedback event may consist of a single or of multiple signals (see section 3.2). The length of a feedback event is defined by the start of the first signal involved and the end of the last signal. To separate one feedback event from the next feedback event, we determined two options. Either, two subsequent feedback events were separated by 300ms<sup>10</sup> of no movements or talk related to giving feedback. Or, two different feedback events could occur consecutively, distinguished by a noticeable change

<sup>10</sup> The number of 300ms was chosen as Trujillo et al. (2018; 2019) found it to be the approximate minimum length of time that naïve observers need to consistently identify a cessation of movement.

in movement, shape, or direction—for instance, a head nod transitioning into a head shake. We then added annotations on separate tiers for each articulator involved as described in **Table 3**. A key distinction from the earlier modality-agnostic approach by Hodge et al. (2023) lies in our grouping of signs, words, vocalizations (e.g., *mhm*), and mouthings under a single category: *talk* (see Section 3.2). In addition, we introduce an annotation tier labeled feedback type, which specifies the design of each feedback event based on the articulators involved. This includes whether the feedback consists of non-manual signals only, *talk* only, manual gestures only, or a combination thereof. **Figure 3** above shows a still from ELAN exemplifying our annotations and **Table 3** gives an overview of the annotation for various articulators.

Articulators	Description
head	various head movements (e.g., nods, shakes, tilts)
eyebrows	eyebrow movements (raises, frowns)
eyes	eye behaviors (e.g., eyes squinted or widened)
nose	nose-related actions (esp. wrinkling)
cheeks	cheek movements (e.g., puffing)
mouth gesture	mouth actions besides mouthing (e.g., pursing lips, smiles)
shoulders	shoulder movements (e.g., shrugs)
manual gesture	hand/arm gestures (e.g., palm-up)
<i>talk</i> category	mouthings, signs, spoken words, and vocalizations

**Table 3:** Overview of annotated articulators.

During the annotation process, we continuously refined the annotation scheme, incorporating additional variables and annotations. For instance, we reannotated smiles and laughter according to the Smiling Intensity Scale (Gironzetti et al. 2016). In this way, our annotation scheme is grounded both in the previous literature as well as in our data. As our coding scheme developed, we also performed various iterations of coding for all data. In addition, we did various rounds of corrections. In this way, most feedback events have been annotated by at least two annotators. Our coding scheme where all the tier values are explained in detail can be found in the Appendix. In total, we identified around 1,900 feedback events in our data, comprising roughly 3,500 feedback signals.

We faced challenges in annotating certain multimodal features; in particular, (mutual) eye gaze was excluded from the analysis. Annotating eye gaze in video data using ELAN proved difficult, leading to inconsistent results and hindering the integration of gaze into the analysis. To address this gap, future research will employ eye-tracking technology, which we plan to use to improve the accuracy of gaze annotations. Apart from the challenges with gaze, we did not

encounter noticeable difficulties in identifying non-manual signals in our datasets, even though some dyads in these corpora (e.g., in the Russian and RSL data) were filmed from a more lateral camera angle than those in the DGS Corpus. Importantly, this did not result in a higher number of unclear annotation values for these dyads compared to the other languages.

In order to assess the consistency of the annotations across coders, we calculated inter-annotator agreement on the articulator most frequently involved in feedback—the head. To this end, we re-annotated a randomly selected subset of the data comprising roughly 50% of all annotated head movements (843 out of 1603) drawn from all languages. These items had not been previously annotated by the respective coder. The onsets, offsets, and durations of the annotations were pre-annotated by the authors and Deaf and hearing assistants and therefore held constant; only the head movement type and the feedback event type were annotated. The resulting two annotation sets were then compared for inter-rater agreement. We calculated Fleiss’ generalized kappa in R version 4.5.1 (R Core Team 2025) (unweighted, 0.95 confidence level) with the function `fleiss.kappa.raw()` from the package `irrCAC` (Gwet 2019). The values show an inter-annotator agreement much higher than chance agreement. With 0.72, the kappa coefficient indicates substantial agreement between the two annotators (Landis & Koch 1977: 165). While these inter-coder reliability calculations concern only the content of the annotations, not their temporal location or duration, the fact that all annotation procedures involved subsequent corrections provides additional assurance that the annotations reflect a reasonable degree of coder consensus. A further limitation that needs to be acknowledged is that the agreement score pertains exclusively to the head tier and does not include less frequent signals—such as eyebrow or mouth movements—which are known to be more challenging for annotators to classify consistently (Esselink et al. 2024).<sup>11</sup>

For the *talk* category, we classified the transcribed elements into subcategories based on their function. These categories are summarized in **Table 4**.

CATEGORY	INCLUDED ELEMENTS
yes-like elements	Equivalents of ‘yes’ in sign, mouthing and speech; nasal response tokens like <i>mhm</i> , typically used in continuer and acknowledgement functions
ah-like elements	Mouthing and speech <i>ah</i> , other change-of-state elements like German <i>ach so</i> ‘ah ok’, typically used in newsmark function
Assessing elements	Evaluative adjectives like German <i>interessant</i> ‘interesting’
Other	Elements that did not fit the other categories

**Table 4:** Summary of subcategories of the *talk* category.

<sup>11</sup> Of course, manual annotations are never fully objective. However, mistakes occur in automatic annotations as well, which is why we are confident that our data offer a good representation of the design of feedback events in our corpora.

### 4.3 Analysis

As our approach is completely exploratory, our analyses do not involve any hypothesis testing. All analyses were performed in R (R Core Team 2025). To investigate the variability between feedback producers of the same language and compare it with cross-linguistic differences, we created a heatmap dendrogram in order to expose the relative association between a feedback producer and an articulator. Our inspiration for employing this method for comparing signers and speakers stems from Hodge et al. (2023). The heatmap dendrogram in **Figure 6** and the heatmaps in **Figures 8** and **9** were created with the function `pheatmap()` from the package `pheatmap` (Kolde 2019). In the heatmap dendrogram, hierarchical clustering is employed in order to identify clusters among the signers and speakers in our sample, as well as the annotated articulators. We did not scale the data in the heatmaps, as we aim to show the overall frequency of a certain articulator in the composition of feedback events of a certain interactant on the basis of the percentage of feedback events that contain that particular articulator. The ideal number of clusters for the data used in the heatmap dendrogram was calculated with the function `NbClust()`<sup>12</sup> from the package `NbClust` (Charrad et al. 2014). The other plots were created with `ggplot2` (Wickham 2016) and, in part, the `GGally` package (Schloerke et al. 2024).

## 5 Results

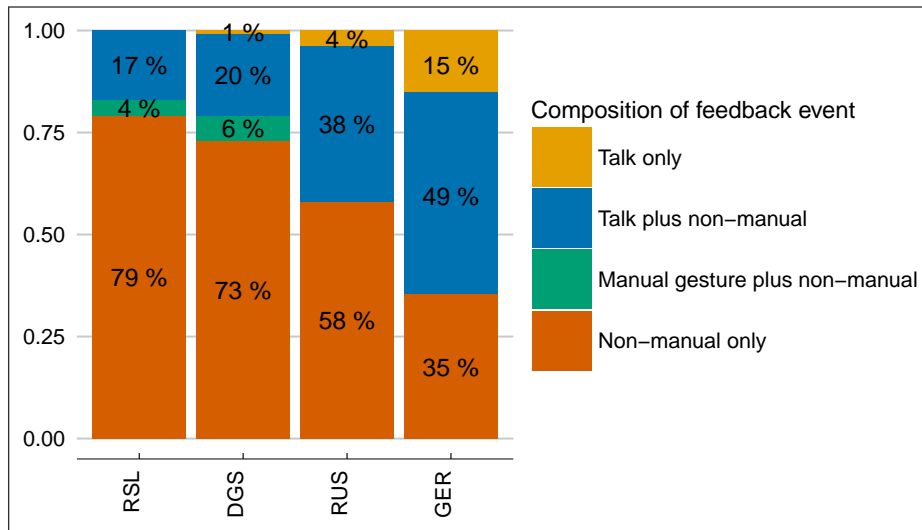
### 5.1 Head nods constitute the most frequent feedback signal in both sign and spoken languages

Across modalities, all four languages show only small percentages of feedback events without any non-manual elements. This is visualized in **Figure 4**, showing for each language the percentages of different combinations of types of feedback signals. A large majority of feedback events in all four languages is constituted by or contains at least one non-manual signal. As some of the numbers are too small to be visible in the figure, we also provide the absolute and relative frequencies of the different feedback event configurations in **Table 5**.

FEEDBACK COMPOSITION	DGS	RSL	GER	RUS
Talk only	5 (0.85%)	1 (0.3%)	81 (15%)	16 (4%)
Manual gesture only	1 (0.15%)	0	0	0
Talk plus non-manual	118 (20%)	68 (17%)	258 (49%)	158 (38%)
Manual gesture plus non-manual	35 (6%)	16 (4%)	1 (0.2%)	1 (0.2%)
Non-manual only	426 (73%)	312 (79%)	185 (35%)	244 (58%)

**Table 5:** Frequencies of feedback event compositions across languages.

<sup>12</sup> We used the following settings: `distance = "euclidean"`, `method = "ward.D2"`.



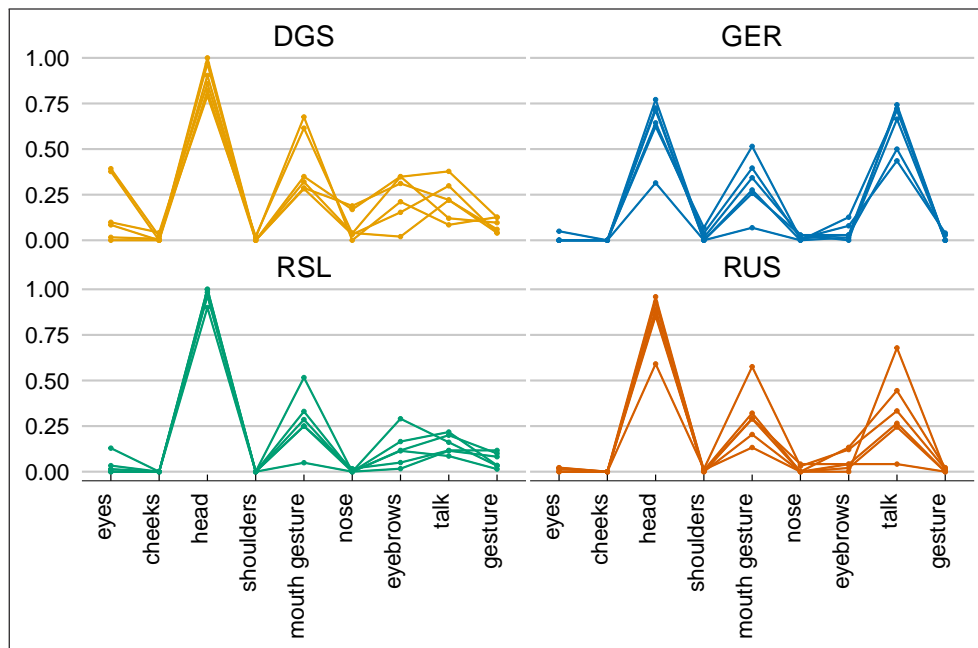
**Figure 4:** Feedback events across languages: Most feedback events consist of or comprise a non-manual element.

**Figure 4** also shows a higher frequency of *talk* for the spoken languages in comparison to the sign languages, particularly in German. However, most feedback events that contain *talk* also contain a non-manual element. We can also observe that feedback events consisting of *talk* alone are virtually absent in the two sign languages.<sup>13</sup> In the two spoken languages, they exist but are relatively rare, German showing the highest proportion with roughly 15%.

In our data, we furthermore observe that the head is the most pervasively used articulator in all four languages. In **Figure 5**, the relative frequencies of articulators employed in feedback events are visualized. Each panel shows a language, with lines representing individual signers/speakers. On the x-axis, the different articulators are placed. The y-axis captures the proportion of feedback events containing the articulator.

**Figure 5** shows that in all languages, the head is the most employed articulator. For the spoken languages, the second most employed articulator is *talk*, followed by mouth gestures. For the sign languages, in contrast, mouth gestures are more pervasively used than *talk*. Moreover, in the two sign languages we can observe a high variability in the employment of *talk*, eyebrow, and nose signals. The use of manual gestures and eyes is somewhat more pervasive in sign than in spoken languages, but generally low. Cheeks and shoulders are only very seldomly mobilized to formulate feedback events across the four languages. In sum, the relative rankings of articulators in **Figure 5** suggest that we are dealing with quantitative rather than fully qualitative differences between sign and spoken languages.

<sup>13</sup> This is consistent with findings for responses to assertions in DGS where non-manuals are also pervasive (Loos et al. 2024: 445).



**Figure 5:** Coordinate plot of articulators involved in feedback: Head is the most frequent articulator in all languages. Each line represents an individual interactant.

Regarding the actual shape of the head movements, these are also relatively similar across languages. In **Table 6**, the frequencies of multiple and single nods as well as other head movements are summarized. It can be observed that across languages, multiple nods constitute the most frequent type of head movement. Taken together, multiple and single nods account for the largest part of head movements during feedback across languages.

LANG.	MULTIPLE NODS	SINGLE NOD	OTHER HEAD MOVEMENTS	TOTAL
DGS	313 (61%)	53 (10%)	147 (29%)	513
RSL	251 (65%)	79 (21%)	55 (14%)	385
GER	183 (52%)	94 (27%)	73 (21%)	350
RUS	249 (70%)	61 (17%)	45 (13%)	355

**Table 6:** Frequencies of different head movements across languages.

In order to investigate the similarity of feedback event configurations across languages, **Table 7** shows some of the most frequent signal combinations from our data and their frequencies in the four languages. Multiple head nods without any other additional signal constitute the most frequently employed feedback event configuration in all four languages. In the spoken languages, this is followed by multiple head nods combined with a *yes*-like element (e.g., equivalents of *yes*

or *mhm*). This combination does not play such a large role in the two sign languages, which, in contrast, have as their second most frequent configuration a single nod without any further signal. A further combination that is employed in all languages with some frequency is multiple head nods combined with a closed mouth smile. Regarding the use of a *yes-like* talk element without further signals, only speakers of spoken German reach a proportion above 10%. In sum, a head nod is the most pervasive head movement during feedback across languages (**Table 6**). *Talk*, e.g. in the form of a *yes-like* talk element, plays a more important role in the spoken than in the sign languages.

FEEDBACK EVENT DESIGN	DGS	RSL	GER	RUS	TOTAL
Multiple head nods	111 (19%)	145 (37%)	77 (15%)	118 (28%)	451 (23%)
Multiple head nods combined with <i>yes-like</i> talk element	16 (3%)	16 (4%)	75 (14%)	74 (18%)	181 (9%)
Single head nod	24 (4%)	55 (14%)	29 (6%)	28 (7%)	136 (7%)
<i>yes-like</i> talk element	2 (<1%)	0	67 (13%)	15 (4%)	84 (4%)
Multiple head nods with closed mouth smile	14 (2%)	23 (6%)	12 (2%)	22 (5%)	71 (4%)
Single nod combined with <i>yes-like</i> talk element	3 (<1%)	1 (<1%)	43 (8%)	17 (4%)	64 (3%)
Total feedback events	585	397	525	419	1926

**Table 7:** Most frequent signal combinations across languages.

While a more detailed investigation of the exact shapes of feedback events is left for future research, **Table 8** offers a glimpse of the frequencies of some of the most frequent non-manual signals (excluding head movements). The table lists absolute frequencies and the percentage of feedback events that contain the signal (alone or combined with other signals) per language.

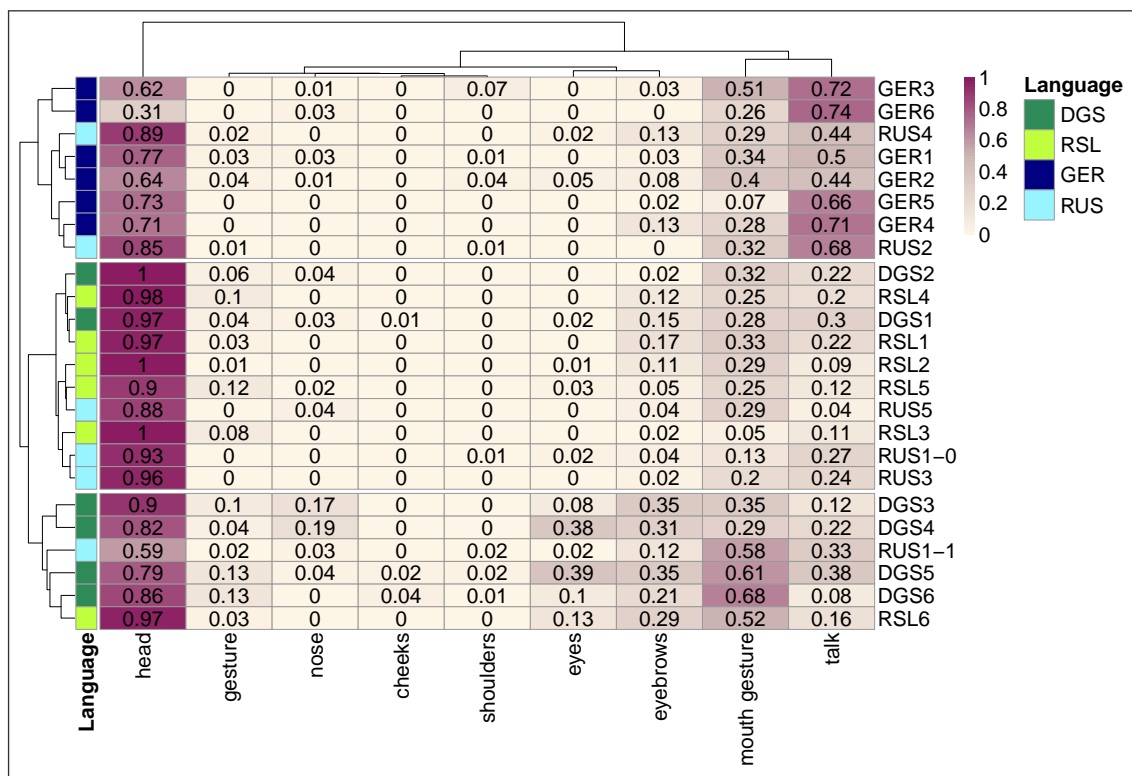
SIGNAL	DGS	RSL	GER	RUS
Eyebrow raise	107 (18%)	44 (11%)	20 (4%)	15 (4%)
Closed mouth smile	72 (12%)	38 (10%)	43 (8%)	70 (17%)
Laugh	9 (1.5%)	9 (2%)	35 (7%)	20 (5%)
Nose wrinkle	48 (8%)	1 (<0.5%)	3 (0.5%)	2 (0.5%)
TOTAL	585	397	525	419

**Table 8:** Frequencies of some of the most frequent non-manual signals, excluding head movements.

## 5.2 Signers and speakers employ a range of feedback styles

To tackle the variability between signers/speakers of the same language that already becomes apparent in **Figure 5**, we created a heatmap dendrogram (**Figure 6**). The numbers in the cells indicate the proportion of feedback events that contain a signal from the respective articulator for the particular interactant. This includes both signals produced on their own or in combination with other signals. The dark color stands for a high percentage of use of a certain articulator by a given interactant, while the light color indicates a low percentage. For instance, the speaker GER3 represented by the first line in the graph employs head movements in 62% of all feedback events she produces. 51% of her feedback events contain a mouth gesture, and 72% contain *talk*.

Moreover, the heatmap dendrogram reveals how similar the different articulators and the interlocutors are to each other. This allows us to investigate whether feedback producers of the same language, the same modality, or the same cultural background will cluster together. Based on the calculation of the ideal number of clusters (see Section 4.3), the interactants in our sample form three basic clusters. These three clusters are separated from each other in the graph for better visibility.



**Figure 6:** Heatmap dendrogram: Interactants form three clusters representing three different feedback styles.

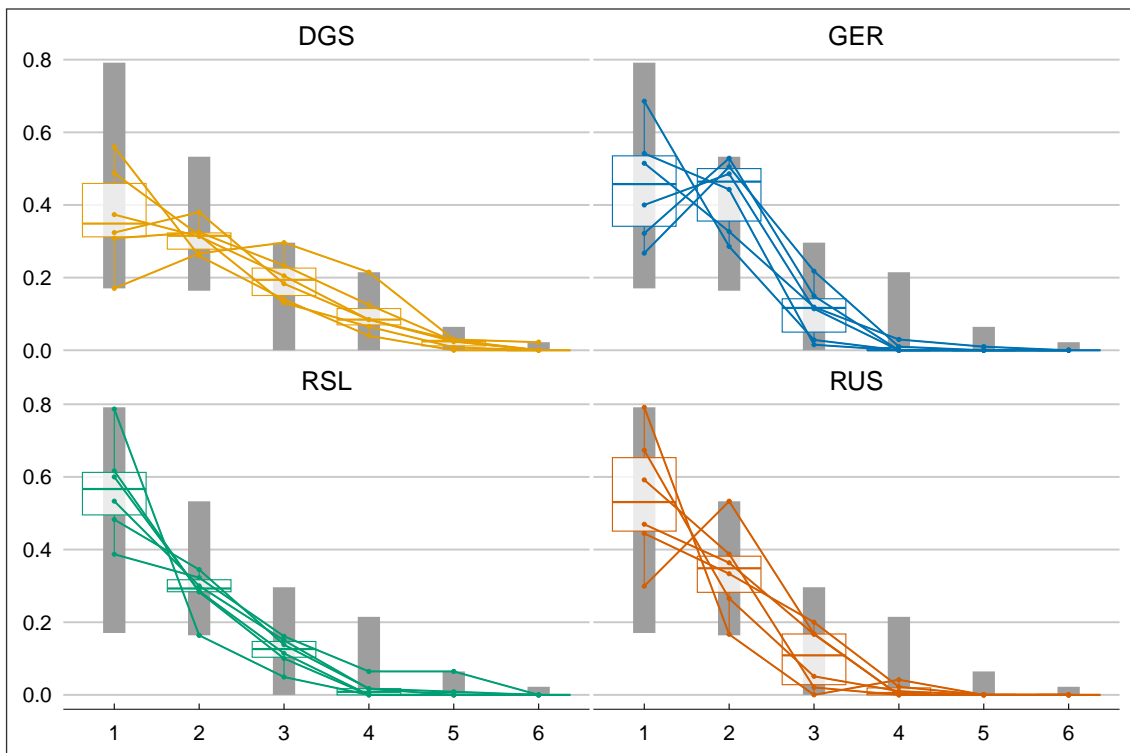
The heatmap dendrogram in **Figure 6** echoes our findings from **Figure 5** above: The head is the most pervasively employed articulator, followed by mouth gesture and *talk*. These articulators also play a role in distinguishing the three largest clusters of interactants in the data: There is one group of interactants (containing all speakers of German and two speakers of Russian) that mostly employ these three articulators. These speakers show a higher proportion of feedback events containing *talk*, and many of them somewhat lower values for head. The second group, containing five signers of RSL, two signers of DGS, and three speakers of Russian, is defined by particularly high percentages of head movements, and comparatively low percentages for all other articulators. The third group, featuring four signers of DGS, one signer of RSL, and one speaker of Russian is characterized by the employment of a more variable and broader range of non-manual articulators, showing higher values for mouth gesture, eyebrows, and eyes. We propose that the clusters in **Figure 6** manifest three different feedback styles: A style that relies more on *talk* and less on head movements; a style that relies mostly on head movements; and a style that employs a broader range of non-manual articulators (for discussion see Section 6.2.).

**Figure 6** shows a clustering according to language, which is however only partial and fuzzy. The speakers of German cluster together due to their higher frequency of the *talk* articulator, whereas most signers of DGS cluster together due to their higher rates of mouth gesture, eyebrows and eyes. The middle group, however, defined by its reliance on very high percentages of head movements, is composed of most RSL signers, but also contains speakers of Russian and signers of DGS. This shows that, while on average speakers of the spoken languages tend to rely on *talk* more than the sign languages, and some signers (mostly those of DGS) tend to rely on a broader variety of non-manual articulators, these two possibilities constitute two extremes on a scale of reliance on *talk* and visual multi-articulator expression. In between, we find signers and speakers who rely on head movements to a large extent. Moreover, in **Figure 6** we can observe that cultural background does not seem to play a major role in the clustering, as there is a complete separation between DGS signers and German speakers. In all three clusters, we find interactants from both cultural backgrounds.

These observations are also reflected in the number of articulators employed in feedback events. In **Figure 7**,<sup>14</sup> we can observe that signers of DGS show higher percentages of feedback events with three or four articulators involved (and consequently less with one articulator only). This fits very well with the observation that most DGS signers form part of the group employing the multi-articulator style in **Figure 6**. German speakers, in contrast, show a relatively high percentage of feedback events with two articulators, which fits the high rates of *talk* while still maintaining head as the most important articulator for those speakers in **Figure 6**. Moreover, we examined all feedback events in our data sample to determine how many consisted of a single

---

<sup>14</sup> The gray bars indicate the spread across all languages and interactants. Each line represents an individual.



**Figure 7:** Proportions of feedback events composed of one to six signals across the four languages. Single- vs. multiple-signal events counts are as follows: DGS (202 / 383), GER (230 / 295), RSL (227 / 170), RUS (206 / 213). In total, 865 events contain a single signal and 1,061 contain multiple signals.

signal versus multiple signals. Overall, multiple-signal events ( $n = 1,061$ ) occur more frequently than single-signal events ( $n = 865$ ), which supports our holistic approach to feedback.<sup>15</sup>

Furthermore, our data give a first hint at intra-speaker variability. The Russian speaker (RUS-1) who features in two of our recordings—one with a previously unknown person (where RUS1 is indicated as RUS1-0) and one with a person she knows (RUS1-1)—is assigned to two different groups, which means that she employs different feedback styles in the two conversations. In the conversation with the person she knows, she is in the group employing a broader range of non-manual signals, basically due to the higher percentage of mouth gestures. In the conversation with the stranger, she is in the group mostly relying on head movements, showing a much larger percentage of head movements (0.93 vs. 0.59). This intra-individual difference cannot be explained by interpersonal alignment (Rasenberg et al. 2020), as the two Russian speakers with whom she interacts are both in the group relying more on talk (RUS2 and RUS4). This finding suggests that it will be promising to put a focus on intra-individual variation in the future, paying attention to feedback styles in different situations.

<sup>15</sup> Some language-specific differences are visible and should be addressed in future research.

Lastly, different types of feedback signals are clearly associated with different feedback functions, which is very likely to in turn contribute to the emergence of the feedback styles we observe.<sup>16</sup> If an interlocutor follows a narration by the other interactant providing mostly continuers as feedback, their feedback style will differ from a situation where the same interactant is participating in a very involved way, responding with higher amounts of assessments that offer the interactant's subjective evaluation, or of newsmarks that index the remarkability of information (Marmorstein & Szczepek Reed 2023). In the latter case, we can expect higher proportions of facial expressions such as smiles or laughter (assessments) or eyebrow raises (newsmarks). Feedback styles, then, are partly shaped by feedback functions that feature prominently in the particular interaction. Future research will therefore examine feedback functions and their correlations with both the feedback signals and the feedback styles proposed here. In addition, we aim to control for the content of each conversation in order to more fully disentangle the influence of feedback function from other factors influencing the distributional patterns observed, such as community-wide conventions and personal preferences.

## 6 Discussion and theoretical implications

### 6.1 The multimodal and multi-channel nature of feedback

In this paper, we compare the formulation of feedback events in four languages—two signed and two spoken—while controlling for cultural background. Our data replicate earlier findings, showing that feedback exhibits some variability across languages, individuals, and can also vary for a single individual according to the situation.

In addition, our data suggest that language modality does play a role in the relative ranking of the different articulators in the composition of feedback events. While *talk* (i.e., manual signs and mouthing) is available to and employed by signers to formulate feedback events, it is used less pervasively than *talk* (i.e., spoken words and vocalizations) in spoken languages. However, the head emerges as the most pervasively used articulator for formulating feedback events across all languages in our sample, with multiple head nods without any accompanying signals constituting the most frequent feedback configuration. Our findings are thus consistent with previous findings on conversational feedback and the proposal of a shared conversational infrastructure for feedback and social interaction in general (Lutzenberger et al. 2024).

This infrastructure is inherently multimodal and multi-channel, involving multiple articulators in both sign and spoken languages.

These findings warrant an explanation. What are the advantages of a multimodal and multi-channel system for communication? We would like to suggest that an inherently multimodal and multi-channelled infrastructure for feedback (and conversation in general) allows signers and

---

<sup>16</sup> We are grateful to an anonymous reviewer for bringing up this point.

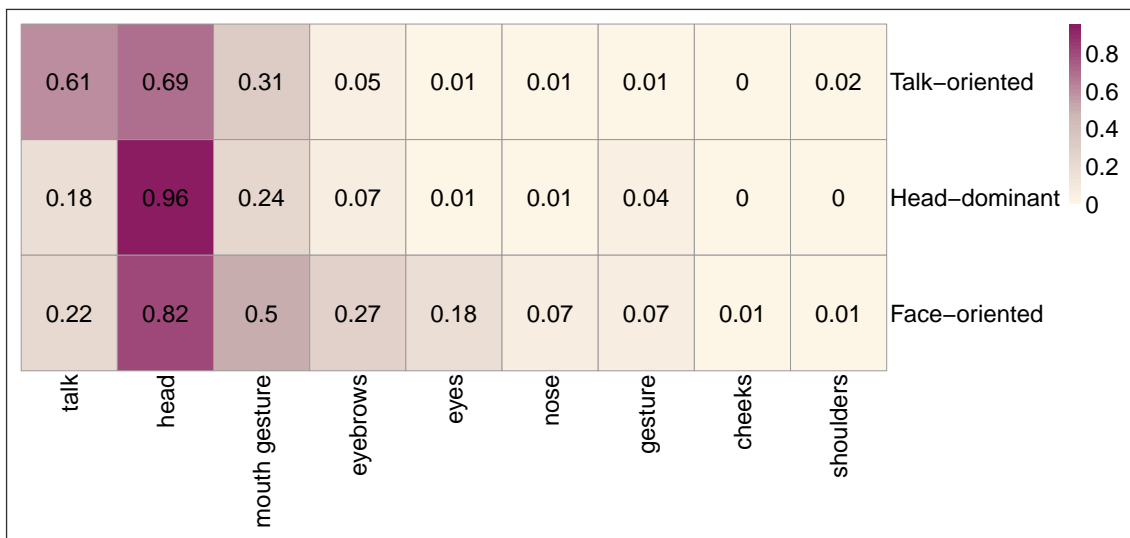
speakers to solve at least three conversational problems. First, the use of head movements and/or other visual signals allows for providing feedback without intruding upon the interlocutor's turn, in line with earlier proposals (Dingemanse et al. 2022; Börstell 2024). Small multiple head nods can be produced in overlap with an ongoing turn without interrupting the current signer or speaker (whereas a manual gesture might be interpreted as an attempt to take the floor) (see Bauer et al. (2024) for kinematic properties of feedback head nods). Second, the possibility of producing visual signals in overlap with the interlocutor's turn also comes with the opportunity of early signaling which action a signer or speaker is intending to perform in the upcoming turn, which may help the interlocutor to identify that action more quickly (Holler 2025). Third, a multimodal infrastructure for feedback allows interactants to evade linearity and thus flexibly express different meanings, including potentially the expression of different interactional functions at the same time (e.g., in the form of a head nod functioning as a continuer combined with a signed or spoken lexical assessment). A single-channel infrastructure, in contrast, implies a linear delivery, allowing for meanings to be expressed only consecutively rather than in parallel. The multimodal and multi-channel infrastructure for conversation thus provides signers and speakers with a system that is much more flexible than a single-channel infrastructure could be.

In addition to offering solutions to these three conversational problems, we suggest that the multimodal infrastructure also raises the question of who is the beneficiary of feedback. What is usually emphasized is the function of feedback of offering interpretable information to the current signer/speaker. Under this account, the recipient of the feedback is the sole or main beneficiary. But the multimodal infrastructure questions this position, as head nods may not only indicate to the addressee that the producer of the nod is still attentive to their talk. Rather, we suggest that it may potentially afford processing benefits for the person who produces the nod. Furthermore, while it is clear that facial gestures in conversation are not mere emotional expressions but rather serve pragmatic and interactional functions (Bavelas & Chovil 2018), this does by no means imply that they may not in addition constitute expressions of emotion that allow the speaker to regulate their emotional state. This is particular relevant in feedback, as feedback is produced *in reaction* to the interlocutor's turn. Here, emotional reactions may coincide with the pragmatic function to be conveyed, e.g., a state of surprise leading to a reaction in the form of raised eyebrows may coincide with the the pragmatic function of indicating that the information provided by the interlocutor is conceived as surprising. For manual gestures, it is well established that they are also produced when they cannot be seen by the interlocutor (Bavelas et al. 1992; Iverson & Goldin-Meadow 1998; Mol et al. 2011). This suggests that their production may potentially not only serve the addressee, but could also impact the cognitive processes of the producer (Goldin-Meadow & Beilock 2010). However, the potential role of non-manual gestures such as head nods in altering the producer's cognitive processes has yet to be

explored in future investigations (Mori et al. 2022). Our data strongly suggest that this will be a worthwhile path.

## 6.2 Modelling multimodal feedback styles

Based on our analysis using the heatmap dendrogram (Figure 6), we identified three distinct feedback styles in our data. These styles differ in the relative prominence of various articulators used during feedback events. These findings are compatible with a shared interactional infrastructure for feedback among sign and spoken languages (Lutzenberger et al. 2024), where, however, interactants can choose among different feedback styles. These styles, we argue, can be ordered on a scale based on the pervasiveness of the different signals, as shown in the heatmap in Figure 8. This heatmap offers a summary of the three feedback styles identified, where each tile in the heatmap shows, per channel, the mean of the proportions for all interactants that were classified to belong to that style in Figure 6. Dark tiles indicate a high mean proportion, light-color tiles a low mean proportion. We chose to build this figure with the mean proportion in order to provide some potentially generalizable observations on the three feedback styles.

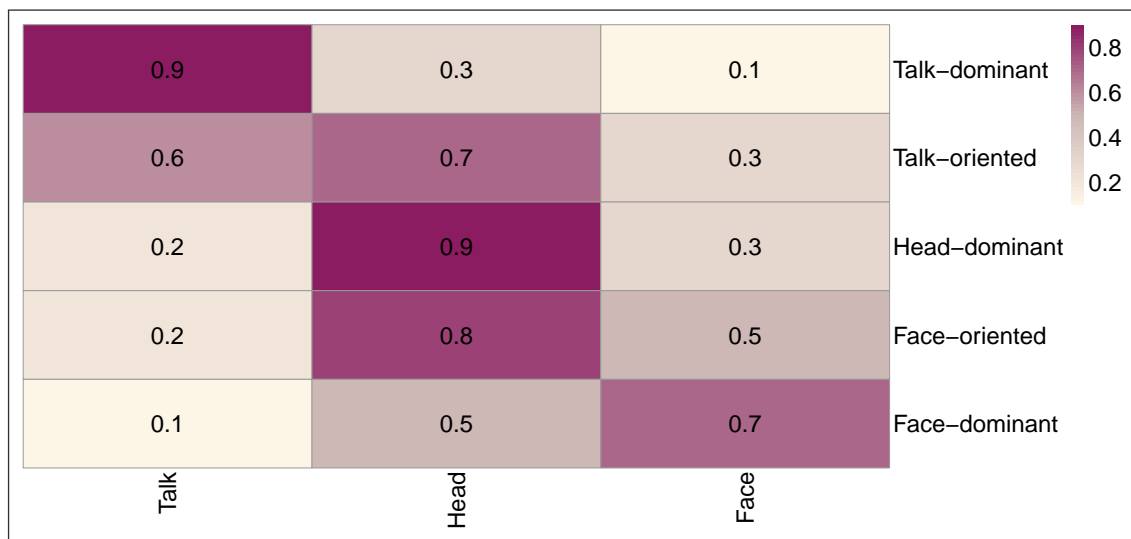


**Figure 8:** Heatmap of feedback styles identified with mean articulator proportions.

The mean proportions in Figure 8 suggest that feedback styles may form a gradient pivoting around the head articulator. In the Head-dominant style in the middle, the head is the most prominent articulator. In the Talk-oriented style, talk becomes more frequent, whereas the proportion of head movements drops. In the Face-oriented style, finally, other non-manual articulators are used more often, while head also drops somewhat. This suggests that when articulators other than the head become more pervasive in the feedback of an interactant, head

movements become less prominent. We propose that from this observation we can derive a model that provides testable predictions for future research on feedback styles.

In **Figure 9**, we visualize the proposed model and summarize its predictions regarding the distribution of articulators in feedback styles. The three styles in the middle represent the styles we discovered in this paper. The means are extrapolated from **Figure 8**. As the proportion of feedback events containing head movements must decrease when other articulators (such as talk or non-manual gestures) become more prominent, the model predicts the theoretical existence of two additional feedback styles: one that is Talk-dominant, and another that is Face-dominant. These two styles represent the hypothetical endpoints of the model’s continuum.



**Figure 9:** Heatmap visualizing the model with observed and predicted feedback styles.

Importantly, unlike the three empirically observed styles in our data, these endpoint styles are theoretical projections—they are not based on actual measured data but are included to illustrate the model’s full conceptual range. To our knowledge, such extreme styles have not yet been explicitly described in the literature, although some preliminary observations (e.g., on Yurakaré and Yélf Dnye<sup>17</sup>) hint at their potential existence.

The model also predicts that no viable feedback style should exhibit simultaneously high proportions of all three articulator types (i.e., talk, head, and other non-manuals). We speculate

<sup>17</sup> A preliminary investigation of a conversational corpus (van Gijn et al. 2011) of Yurakaré (isolate, Bolivia) suggests that in this language, head movements are extremely rarely used in feedback, with speakers relying mostly on spoken elements, thus potentially representing the Talk-dominant style. The Papuan language Yélf Dnye, where eye blinks and eyebrow flashes are regularly employed as continuers (Levinson 2015: 406), may be a candidate for the Face-dominant style.

that such an all-dominant configuration would likely be inefficient or cognitively taxing for both the producer and the addressee to process, as activating and integrating signals across all available channels might overload the interactional system rather than enhance it. This suggests that efficient feedback systems may rely on channel selection and modulation, rather than maximal channel activation.

For testing these model predictions and the adequacy of the model for new data, it will be crucial to conduct further studies with pairs of sign and spoken languages matched for cultural background. In addition, it will be relevant to investigate languages in non-(Indo-)European contexts in order to arrive at a more diverse picture.

The conceptualization of the feedback space in terms of a multi-dimensional gradient as visualized in **Figure 9** has another advantage. In addition to an analysis of inter-individual variation as in **Figure 6**, it allows for an investigation and theorization of intra-individual variation. It is quite probable that signers and speakers are capable to adjust their feedback style according to the type and topic of the conversation, their interlocutor, degree of familiarity and thus the extent of shared common ground and even their personal physical and mental state at the time of the conversation. As another testable hypothesis, we propose that signers or speakers will vary along these lines, and that their styles will correspond to those we found in this paper, summarized in **Figure 9**. Of course, it is well possible that other styles will be discovered once more languages—including non-(Indo-) European ones—are investigated and larger samples with more interactants form the basis of our investigations, hopefully made possible through automatized annotations.

### **6.3 Towards a multimodal and interactional theory of the Language Faculty**

Our results furthermore have important implications for the conceptualization of conversational concepts, as well as for theories of the Language Faculty. Specifically, they demonstrate the dominant role of non-manual signals—such as head movements, facial gestures, and other visual signals—in the production of feedback. This challenges the prevailing emphasis on verbal feedback in accounts of interactional phenomena, which have largely neglected the visual dimension of communication.

Recent research, making use of advances in recording and experimental methodologies, has significantly deepened our understanding of co-present interaction, revealing the fundamentally multimodal nature of human communication (Gregori et al. 2023; Henlein et al. 2024). While linguists increasingly recognize the multimodal foundation of language and interaction (Perniss 2018; Holler & Levinson 2019; Özyürek 2021; Rasenberg et al. 2022; Hamilton & Holler 2023; Kendrick et al. 2023; Sandler 2024), fundamental theoretical concepts, such as conversational turns and interactional feedback mechanisms, are often theorized in unimodal terms, leaving aside their inherently multimodal character.

We propose a shift in this perspective: these concepts should be theorized as multimodal from the outset as also currently suggested by Holler (2025) for ‘social action’. For instance, if a hand gesture precedes speech, why should the conversational turn be defined as beginning with the onset of the first syllable?

We argue that the boundaries of conversational turns—both their beginnings and endings—cannot be defined solely by spoken or signed items (see also Kendrick et al. 2023). Manual and non-manual gestures should instead be regarded as intrinsic components of conversational turns, as we have demonstrated for feedback in this study. A single head nod or head tilt, alone or in combination with other non-manual signals (e.g., widened eyes, eyebrow movements) or with manual gestures (e.g., palm-up gestures), constitutes an integral element of a feedback event.

Our results furthermore reinforce the urgent need for models of language and the Language Faculty that engage with the inherently multimodal nature of human communication. The Multimodal Language Faculty (MLF) model, a cognitive framework recently developed by Cohn & Schilperoord (2024), aims to account for both unimodal and multimodal language use, as well as other forms of communication across various modalities. While this model shows considerable flexibility, particularly through its proposed Multimodal Parallel Architecture, it remains unclear how interaction and dialogue are formally represented, as these components are not explicitly addressed in the current formulation of MLF. Our data highlight the need to extend such models to incorporate fundamentally interactional phenomena such as feedback, where the focus is not primarily on truth-conditional meaning but rather on interactional function. In contrast, the Interactional Spine Hypothesis (Wiltschko 2021) offers a highly detailed theoretical account of interaction, providing new perspectives on several interactional mechanisms including responsive actions. However, this model does not currently provide a framework for incorporating meaningful visual signals as core interactional phenomena. Extending this model to incorporate visual signals will be a fruitful path in the future, as the model explicitly predicts multi-functionality of linguistic items and is thus very well-suited for the integration of multi-functional non-manual signals. Taken together, our results suggest that visual, non-manual signals are an integral component of human interaction. We are therefore in urgent need of theoretical models that integrate both the visual and interactional dimensions of human language, moving beyond speech-centered paradigms.

## **7 Conclusion and future work**

This study explored the composition of feedback events across four languages and two language modalities, employing a novel cross-linguistic and cross-modal approach that considers the full constellation of communicative resources used to provide feedback. The findings reveal that non-manual signals are fundamental to conversational interaction. This has significant implications for linguistic theory, suggesting the need to move beyond purely speech-based

models. A reconceptualization is required to account for the multimodal nature of interaction, as speech alone does not provide a complete picture of how communication is performed.

Our results also point to the importance of examining intra-individual variation in feedback. In future studies, we plan to compare interactions between familiar and unfamiliar conversation partners. By employing eye-tracking devices during data collection, we aim to measure (mutual) eye gaze during, before, and after feedback events, allowing for a more nuanced understanding of the role of gaze behaviors in feedback. The inclusion of eye gaze as a key component of feedback will be a significant focus of our future work, addressing the limitations in this study where annotating gaze behavior manually was challenging.

Additionally, we recognize the absence of a prosodic analysis of verbal feedback in this study, as we focused primarily on articulators. Future research will aim to incorporate prosodic features alongside gaze, offering a more comprehensive understanding of feedback mechanisms.

Expanding the research to include a wider variety of linguistic and cultural contexts could yield valuable insights. For example, in Bulgarian communication, agreement or affirmation is often signaled through a lateral head movement. Investigating whether these gestures influence the use of head movement during feedback interactions would be an interesting area for future research. Moreover, examining contexts where direct gaze is culturally less common, such as among speakers of Tzeltal (Mayan) (Rossano et al. 2009), could broaden our understanding of feedback mechanisms. Studies like these will help to illuminate how different linguistic communities navigate feedback, offering a richer, cross-cultural perspective on multimodal interactional strategies.

Through this research, we contribute to the theoretical framework surrounding multimodal feedback, advancing the understanding of feedback mechanisms within diverse linguistic and interactional contexts. The proposed model of feedback styles is a first attempt to understand how interlocutors navigate social interaction by adjusting their multimodal behaviour. Studies like ours pave the way toward a more comprehensive understanding of how multimodal turns operate across different languages, helping to illuminate the universal and variable aspects of feedback in human communication.

---

## Appendix: Coding scheme for annotation of multimodal feedback events

The following table presents our annotation scheme for multimodal feedback in signed and spoken interactions. While developing this scheme, we drew inspiration from prior literature, incorporating certain labels and abbreviations (e.g., head and mouth gestures: Burkova (2015), <http://rsl.nstu.ru/>; smiles and laughter: Smiling Intensity Scale, Gironzetti et al. (2016); eye blinks: Hömke et al. (2017)) as well as insights from the data analyzed in this study. We excluded body movements and eye blinks from the current analysis and leave the investigation of these two features for future research.

ABBREVIATION	MEANING
<i>Tier mouth gesture</i>	
lbt	biting of the lower lip
ldn	corners of the mouth lowered down
ldr	lips sucked in
lo	lips rounded
lpd	lower lip pushed forward
lpf	lips pushed forward
lp	lips pressed together
lvb	lips tremble
mbl	blowing out air
mo	mouth open
msc	sucking in air
tch	tongue against the cheek
tt	tongue out
cms	closed mouth smile (s1)
oms	open mouth smile (s2)
woms	wide open mouth smile (s3)
lgh	laughing smile or laugh, smiling with jaw dropped (s4)
<i>Tier head</i>	
hnn	many short head nods
sn	small (shallow) head nod
ln	large head nod
lnn	many large nods
mn	mixed nod (e.g. one large nod followed by small nod(s))
hb	head tilt back
hbn	head tilt back with subsequent head nod
hs	head shake

(Contd.)

<b>ABBREVIATION</b>	<b>MEANING</b>
hmb	head move backward
hmf	head move forward
hl	head turn to the left
hlb	head turn to the left & tilted backwards
hlf	head turn to the left & tilted forward
hr	head turn to the right
hrb	head turn to the right & tilted backwards
hrf	head turn to the right & tilted forward
hth	head lowering
hths	head lowering & head shake
ht	head tilted to the right or left shoulder
cu	chin up (no head back tilt)
wig	head wiggle (lateral head movements to both sides, neither shake nor nod)
<i>Tier eyebrows</i>	
bf	eyebrows furrowed (= eyebrows are pulled together)
br	eyebrows raised
brd	eyebrows lowered
<i>Tier eyes</i>	
mbl	multiple eye blinks
sbl	short blink (no longer than 410 ms)
lbl	long blink (longer than 410 ms)
esc	eyes squinted
ew	eyes wide opened
<i>Tier nose</i>	
nw	nose wrinkled
nbl	nose blows out air
<i>Tier cheeks</i>	
chp	cheeks blown out/puffed
chs	cheeks sucked in
<i>Tier shoulders</i>	
shf	shoulders curved forwards
shs	shoulder shrug (raising and lowering of shoulders like "I don't know!")
<i>Tier body</i>	
bb	body leaned backward
bf	body leaned forward
bu	body moves/raises up
bt	body turned to the left/to the right
bl	body leaned to the left/to the right (without turning)

## Data availability

Three of the corpora associated with this article are published and thus accessible, either open access (DGS) or upon request (RUS, RSL). The data set and video examples used in this study as well as the script for data analysis are available at <https://doi.org/10.6084/m9.figshare.30738701>.

## Ethics and consent

We are using available corpora in this study. We only use corpora where informed consent was gathered from all participants. The identities of all participants have been anonymized.

## Funding information

This research was funded by the University of Cologne Excellent Research Program, Funding line Cluster Development Program, project Language Challenges.

## Acknowledgements

We are grateful to the signers and speakers who participated in the corpus data collections. Moreover, we would like to express our gratitude to three anonymous reviewers for their thoughtful and inspiring comments on an earlier version of this paper.

We thank Roman Poryadin, Undine Kuhlmann, Milena Pielen and Lina Herrmann for their assistance with annotations. We also thank our colleagues, in particular Birgit Hellwig, Pamela Perniss, Nikolaus P. Himmelmann and Alice Mitchell for their insightful comments on an earlier version of this research. We are grateful to Klaus von Heusinger for suggesting the term ‘feedback event’.

## Competing interests

The authors have no competing interests to declare.

## Authors' contributions

Conceptualization: AB, SG; Data collection: AB; Basic data annotation: AB, SG, JH, TAH; Detailed data annotation (correction, elaboration): AB, JH, SG; Supervision of student assistants: AB, SG; Data wrangling and statistical analysis: SG; Investigation: AB, SG; Writing: AB, SG. AB and SG contributed equally to this work as joint first authors. The remaining authors are listed in alphabetical order by surname.

---

## References

- Abner, Natasha & Cooperrider, Kensy & Goldin-Meadow, Susan. 2015. Gesture for linguists: A handy primer. *Language and Linguistics Compass* 9(11). 437–451. DOI: <https://doi.org/10.1111/lnc3.12168>
- Allwood, Jens & Cerrato, Loredana. 2003. A study of gestural feedback expressions. In *First Nordic Symposium on Multimodal Communication*, 7–22. Copenhagen: Gothenburg University Publications.
- Allwood, Jens & Cerrato, Loredana & Jokinen, Kristiina & Navarretta, Costanza & Paggio, Patrizia. 2007a. The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Language Resources and Evaluation* 41, 273–287. DOI: <https://doi.org/10.1007/s10579-007-9061-5>
- Allwood, Jens & Kopp, Stefan & Grammer, Karl & Ahlsén, Elisabeth & Oberzaucher, Elisabeth & Koppensteiner, Markus. 2007b. The analysis of embodied communicative feedback in multimodal corpora: A prerequisite for behavior simulation. *Language Resources and Evaluation* 41, 255–272. DOI: <https://doi.org/10.1007/s10579-007-9056-2>
- Allwood, Jens & Nivre, Joakim & Ahlsén, Elisabeth. 1992. On the semantics and pragmatics of linguistic feedback. *Journal of Semantics* 9(1). 1–26. DOI: <https://doi.org/10.1093/jos/9.1.1>
- Ameika, Felix K. & Terkourafi, Marina. 2019. What if...? Imagining non-Western perspectives on pragmatic theory and practice. *Journal of Pragmatics* 145. 72–82. DOI: <https://doi.org/10.1016/j.pragma.2019.04.001>
- Andries, Fien & Meissl, Katharina & Vries, Clarissa de & Feyaerts, Kurt & Oben, Bert & Sambre, Paul & Vermeerbergen, Myriam & Brône, Geert. 2023. Multimodal stance-taking in interaction—A systematic literature review. *Frontiers in Communication* 8. 1187977. DOI: <https://doi.org/10.3389/fcomm.2023.1187977>
- Backer, Charlotte. 1977. Regulators and turn-taking in American Sign Language discourse. In Friedman, Lynn (ed.), *On the other hand: New perspectives on American Sign Language*, 138–139. NY: Academic Press.
- Bauer, Anastasia. 2023. Russian multimodal conversational data. Data Center for the Humanities, University of Cologne. DOI: <https://doi.org/10.18716/DCH/A.00000016>
- Bauer, Anastasia & Kuder, Anna & Schulder, Marc & Schepens, Job. 2024. Phonetic differences between affirmative and feedback head nods in German Sign Language (DGS): A pose estimation study. *PLOS ONE* 19(5). e0304040. DOI: <https://doi.org/10.1371/journal.pone.0304040>
- Bauer, Anastasia & Kyuseva, Masha. 2022. New insights into mouthings: Evidence from a corpus-based study of Russian Sign Language. *Frontiers in Psychology* 12. 779958. DOI: <https://doi.org/10.3389/fpsyg.2021.779958>
- Bauer, Anastasia & Poryadin, Roman. 2023. Russian Sign Language conversations. Data Center for the Humanities, University of Cologne. DOI: <https://doi.org/10.18716/DCH/A.00000028>
- Bavelas, Janet. 1990. Nonverbal and social aspects of discourse in face-to-face interaction. *Text – Interdisciplinary Journal for the Study of Discourse* 10(1–2). 5–8. DOI: <https://doi.org/10.1515/text.1.1990.10.1-2.5>

- Bavelas, Janet & Chovil, Nicole. 2018. Some pragmatic functions of conversational facial gestures. *Gesture* 17(1). 98–127. DOI: <https://doi.org/10.1075/gest.00012.bav>
- Bavelas, Janet & Chovil, Nicole & Lawrie, Douglas & Wade, Allan. 1992. Interactive gestures. *Discourse Processes* 15(4). 469–489. DOI: <https://doi.org/10.1080/01638539209544823>
- Bavelas, Janet & Coates, Linda & Johnson, Trudy. 2000. Listeners as co-narrators. *Journal of Personality and Social Psychology* 79(6). 941–952. DOI: <https://doi.org/10.1037/0022-3514.79.6.941>
- Bavelas, Janet & Coates, Linda & Johnson, Trudy. 2002. Listener responses as a collaborative process: The role of gaze. *Journal of Communication* 52(3). 566–580. DOI: <https://doi.org/10.1111/j.1460-2466.2002.tb02562.x>
- Beach, Wayne A. 1993. Transitional regularities for ‘casual’ “Okay” usages. *Journal of Pragmatics* 19(4). 325–352. DOI: [https://doi.org/10.1016/0378-2166\(93\)90092-4](https://doi.org/10.1016/0378-2166(93)90092-4)
- Bendel Larcher, Sylvia. 2021. *Interaktionsprofil und Persönlichkeit: Eine explorative Studie zum Zusammenhang von sprachlichem Verhalten und Persönlichkeit*. Göttingen: Verlag für Gesprächsforschung.
- Bertrand, Roxane & Ferré, Gaëlle & Blache, Philippe & Espesser, Robert & Rauzy, Stéphane. 2007. Backchannels revisited from a multimodal perspective. *Proc. Auditory-Visual Speech Processing*, 1–5. Hilvarenbeek, Netherlands.
- Bilous, Frances R. & Krauss, Robert M. 1988. Dominance and accommodation in the conversational behaviours of same- and mixed-gender dyads. *Language & Communication* 8(3–4). 183–194. DOI: [https://doi.org/10.1016/0271-5309\(88\)90016-X](https://doi.org/10.1016/0271-5309(88)90016-X)
- Blomsma, Peter & Skantze, Gabriel & Swerts, Marc. 2022. Backchannel behavior influences the perceived personality of human and artificial communication partners. *Frontiers in Artificial Intelligence* 5. 835298. DOI: <https://doi.org/10.3389/frai.2022.835298>
- Blomsma, Peter & Vaitonyté, Julija & Skantze, Gabriel & Swerts, Marc. 2024. Backchannel behavior is idiosyncratic. *Language and Cognition* 16(4). 1–24. DOI: <https://doi.org/10.1017/langcog.2024.1>
- Börstell, Carl. 2024. Finding continuers in Swedish Sign Language. *Linguistics Vanguard* 10(1). 537–548. DOI: <https://doi.org/10.1515/lingvan-2024-0025>
- Boudin, Auriane & Bertrand, Roxane & Rauzy, Stéphane & Ochs, Magalie & Blache, Philippe. 2024. A multimodal model for predicting feedback position and type during conversation. *Speech Communication* 159. 103066. DOI: <https://doi.org/10.1016/j.specom.2024.103066>
- Brunner, Lawrence J. 1979. Smiles can be back channels. *Journal of Personality and Social Psychology* 37(5). 728–734. DOI: <https://doi.org/10.1037/0022-3514.37.5.728>
- Burkova, Svetlana. 2015. Russian Sign Language Corpus. <http://rsl.nstu.ru/>.
- Byun, Kang-Suk & de Vos, Connie & Bradford, Anastasia & Zeshan, Ulrike & Levinson, Stephen. 2018. First encounters: Repair sequences in cross-signing. *Topics in Cognitive Science* 10(2). 314–334. DOI: <https://doi.org/10.1111/tops.12303>

- Cassell, Justine & Thorisson, Kristinn R. 1999. The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence* 13(4–5). 519–538. DOI: <https://doi.org/10.1080/088395199117360>
- Cerrato, Loredana & Skhiri, Mustapha. 2003. A method for the analysis and measurement of communicative head movements in human dialogues. In *Proceedings of AVSP – International conference on audio-visual speech processing*, 251–256. [https://www.isca-archive.org/avsp\\_2003/cerrato03\\_avsp.pdf](https://www.isca-archive.org/avsp_2003/cerrato03_avsp.pdf).
- Charrad, Malika & Ghazzali, Nadia & Boiteau, Véronique & Niknafs, Azam. 2014. NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software* 61. 1–36. DOI: <https://doi.org/10.18637/jss.v061.i06>
- Clancy, Patricia M. & Thompson, Sandra A. & Suzuki, Ryoko & Tao, Hongyin. 1996. The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics* 26(3). 355–387. DOI: [https://doi.org/10.1016/0378-2166\(95\)00036-4](https://doi.org/10.1016/0378-2166(95)00036-4)
- Coates, Jennifer & Sutton-Spence, Rachel. 2001. Turn-taking patterns in Deaf conversation. *Journal of Sociolinguistics* 5(4). 507–529. DOI: <https://doi.org/10.1111/1467-9481.00162>
- Cohn, Neil & Schilperoord, Joost. 2024. *A multimodal language faculty: A cognitive framework for human communication*. London/New York/Dublin: Bloomsbury. DOI: <https://doi.org/10.5040/9781350404861>
- Crasborn, Onno & Sloetjes, Han. 2008. Enhanced ELAN functionality for sign language corpora. In *6th International Conference on Language Resources and Evaluation (LREC 2008) / 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language corpora*, 39–43.
- Dideriksen, Christina & Christiansen, Morten H. & Dingemanse, Mark & Højmark-Bertelsen, Malte & Johansson, Christer & Tylén, Kristian & Fusaroli, Riccardo. 2023. Language-specific constraints on conversation: Evidence from Danish and Norwegian. *Cognitive Science* 47(11). DOI: <https://doi.org/10.1111/cogs.13387>
- Dingemanse, Mark & Enfield, N. J. 2015. Other-initiated repair across languages: Towards a typology of conversational structures. *Open Linguistics* 1(1). 96–118. DOI: <https://doi.org/10.2478/opli-2014-0007>
- Dingemanse, Mark & Liesenfeld, Andreas & Woensdregt, Marieke. 2022. Convergent cultural evolution of continuers (mhmm). In Ravnani, Andrea & Asano, Rie & Valente, Daria & Ferretti, Francesco & Hartmann, Stefan & Hayashi, Misato & Jadoul, Yannick & Martins, Mauricio & Oseki, Yoshei & Rodrigues, Evelina Daniela & Vasileva, Olga & Waciewicz, Slawomir (eds.), *The evolution of language: Proceedings of the joint conference on language evolution (JCoLE)*, 160–167. DOI: <https://doi.org/10.31234/osf.io/65c79>
- Dingemanse, Mark & Roberts, Seán G. & Baranova, Julija & Blythe, Joe & Drew, Paul & Floyd, Simeon & Gisladdottir, Rosa S. & Kendrick, Kobin H. & Levinson, Stephen & Manrique, Elizabeth & Rossi, Giovanni & Enfield, Nick. 2015. Universal principles in the repair of communication problems. *PLoS ONE* 10(9). e0136100. DOI: <https://doi.org/10.1371/journal.pone.0136100>

- Dittmann, Allen T. & Llewellyn, Lynn G. 1968. Relationship between vocalizations and head nods as listener responses. *Journal of Personality and Social Psychology* 9(1). 79–84. DOI: <https://doi.org/10.1037/h0025722>
- Drummond, Kent & Hopper, Robert. 1993. Back channels revisited: Acknowledgment tokens and speakership incipency. *Research on Language & Social Interaction* 26(2). 157–177. DOI: [https://doi.org/10.1207/s15327973rlsi2602\\_3](https://doi.org/10.1207/s15327973rlsi2602_3)
- Duncan, Starkey. 1974. On the structure of speaker–auditor interaction during speaking turns. *Language in Society* 3(2). 161–180. DOI: <https://doi.org/10.1017/S0047404500004322>
- Esselink, L. D. & Oomen, M. & Roelofsen, Floris. 2024. Technical report: Evaluating inter-annotator agreement for non-manual markers in sign languages. DOI: <https://doi.org/10.21942/UVA.25563540.V2>.
- Fenlon, Jordan & Schembri, Adam C. & Sutton-Spence, Rachel. 2013. Turn-taking and backchannel behaviour in British Sign Language conversations. Poster presented at the *11th Theoretical Issues in Sign Language Research Conference*, University College London.
- Fujimoto, Donna T. 2009. Listener responses in interaction: A case for abandoning the term, backchannel. *Bulletin paper of Osaka Jogakuin College* 37. 35–54. <http://ir-lib.wilmina.ac.jp/dspace/bitstream/10775/48/1/03.pdf>.
- Gardner, Rod. 2001. *When listeners talk: Response tokens and listener stance*. Amsterdam/Philadelphia: John Benjamins. DOI: <https://doi.org/10.1075/pbns.92>
- Gipper, Sonja & König, Katharina & Weber, Kathrin. 2023. Structurally similar formats are not functionally equivalent across languages: Requests for reconfirmation in comparative perspective. *Contrastive Pragmatics* 5(1–2). 195–237. DOI: <https://doi.org/10.1163/26660393-bja10097>
- Girard-Groerber, Simone. 2015. The management of turn transition in signed interaction through the lens of overlaps. *Frontiers in Psychology* 6. 741. DOI: <https://doi.org/10.3389/fpsyg.2015.00741>
- Gironzetti, Elisa & Pickering, Lucy & Huang, Meichan & Zhang, Ying & Menjo, Shigehito & Attardo, Salvatore. 2016. Smiling synchronicity and gaze patterns in dyadic humorous conversations. *HUMOR* 29(2). 301–324. DOI: <https://doi.org/10.1515/humor-2016-0005>
- Goldin-Meadow, Susan & Beilock, Sian L. 2010. Action's influence on thought: The case of gesture. *Perspectives on Psychological Science* 5(6). 664–674. DOI: <https://doi.org/10.1177/1745691610388764>
- Goodwin, Charles. 1986. Gestures as a resource for the organization of mutual orientation. *Semiotica* 62(1–2). 29–50. DOI: <https://doi.org/10.1515/semi.1986.62.1-2.29>
- Gregori, Alina & Amici, Federica & Brilmayer, Ingmar & Ćwiek, Aleksandra & Fritzsche, Lennart & Fuchs, Susanne & Henlein, Alexander & Herbort, Oliver & Kügler, Frank & Lemanski, Jens & Liebal, Katja & Lücking, Andy & Mehler, Alexander & Nguyen, Kim Tien & Pouw, Wim & Prieto, Pilar & Rohrer, Patrick Louis & Sánchez-Ramón, Paula G. & Schulte-Rüther, Martin & Schumacher, Petra B. & Schweinberger, Stefan R. & Struckmeier, Volker & Trettenbrein, Patrick C. & Von Eiff, Celina I. 2023. A roadmap for technological innovation in multimodal communication research. In

- Duffy, Vincent G. (ed.), *Digital human modeling and applications in health, safety, ergonomics and risk management*, 402–438. Cham: Springer. DOI: [https://doi.org/10.1007/978-3-031-35748-0\\_30](https://doi.org/10.1007/978-3-031-35748-0_30)
- Gwet, Kilem L. 2019. irrCAC: Computing chance-corrected agreement coefficients (CAC). R-package. <https://cran.r-project.org/web/packages/irrCAC/irrCAC.pdf>
- Hadar, Uri & Steiner, Timothy & Rose, F. Clifford. 1985. Head movement during listening turns in conversation. *Journal of Nonverbal Behavior* 9(4). 214–228. DOI: <https://doi.org/10.1007/BF00986881>
- Hamilton, Antonia F. De C. & Holler, Judith. 2023. Face2face: Advancing the science of social interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 378(1875). 20210470. DOI: <https://doi.org/10.1098/rstb.2021.0470>
- Hanke, Thomas & Schulder, Marc & Konrad, Reiner & Jahn, Elena. 2020. Extending the Public DGS Corpus in Size and Depth. In Efthimiou, Eleni & Fotinea, Stavroula-Evita & Hanke, Thomas & Hochgesang, Julie A. & Kristoffersen, Jette & Mesch, Johanna (eds.), *Proceedings of the LREC2020 9th workshop on the representation and processing of Sign Languages: Sign language resources in the service of the language community, technological challenges and application perspectives*, 75–82. Marseille, France: European Language Resources Association (ELRA). <https://www.sign-lang.uni-hamburg.de/lrec/pub/20016.pdf>.
- Henlein, Alexander & Bauer, Anastasia & Bhattacharjee, Reetu & Ćwiek, Aleksandra & Gregori, Alina & Kügler, Frank & Lemanski, Jens & Lücking, Andy & Mehler, Alexander & Prieto, Pilar & Sánchez-Ramón, Paula G. & Schepens, Job & Schulte-Rüther, Martin & Schweinberger, Stefan R. & Von Eiff, Celina I. 2024. An outlook for AI innovation in multimodal communication research. In Duffy, Vincent G. (ed.), *Digital human modeling and applications in health, safety, ergonomics and risk management*, 182–234. Cham: Springer. DOI: [https://doi.org/10.1007/978-3-031-61066-0\\_13](https://doi.org/10.1007/978-3-031-61066-0_13)
- Heritage, John. 1984. A change-of-state token and aspects of its sequential placement. In Atkinson, J. Maxwell (ed.), *Structures of social action: Studies in Conversation Analysis*, 299–345. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511665868.020>
- Herrmann, Annika. 2020. Prosody: Back-channeling. In Proske, Sina & Herrmann, Annika & Hosemann, Jana & Steinbach, Markus (eds.), *A grammar of German Sign Language (DGS)* (SIGN-HUB Sign Language Grammar Series 71), 1st edn. <https://thesignhub.eu/grammar/dgs?tag=100>.
- Hess, Lucille J. & Johnston, Judith R. 1988. Acquisition of back channel listener responses to adequate messages. *Discourse Processes* 11(3). 319–335. DOI: <https://doi.org/10.1080/01638538809544706>
- Hodge, Gabrielle & Barth, Danielle & Reed, Lauren W. 2023. Auslan and Matukar Panau: A modality-agnostic look at quotatives. *The Social Cognition Parallax Interview Corpus (SCOPIC). Language Documentation & Conservation Special Publication* 12. 85–125. <https://hdl.handle.net/10125/24744>
- Hoffmann, Bettina & Himmelmann, Nikolaus P. 2009. Münster Videokorpus Alltagsgespräche. Unpublished corpus of spoken German.
- Holler, Judith. 2025. Facial clues to conversational intentions. *Trends in Cognitive Sciences* 29(8). 750–762. DOI: <https://doi.org/10.1016/j.tics.2025.03.006>

- Holler, Judith & Levinson, Stephen C. 2019. Multimodal language processing in human communication. *Trends in Cognitive Sciences* 23(8). 639–652. DOI: <https://doi.org/10.1016/j.tics.2019.05.006>
- Hömke, Paul & Holler, Judith & Levinson, Stephen. 2017. Eye blinking as addressee feedback in face-to-face conversation. *Research on Language and Social Interaction* 50(1). 54–70. DOI: <https://doi.org/10.1080/08351813.2017.1262143>
- Iverson, Jana M. & Goldin-Meadow, Susan. 1998. Why people gesture when they speak. *Nature* 396(6708). 228–228. DOI: <https://doi.org/10.1038/24300>
- Jefferson, Gail. 1984. On the organization of laughter in talk about troubles. In Atkinson, J. Maxwell & Heritage, John (eds.), *Structures of social action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press.
- Jefferson, Gail. 1993. Caveat speaker: Preliminary notes on recipient topic-shift implicature. *Research on Language & Social Interaction* 26(1). 1–30. DOI: [https://doi.org/10.1207/s15327973rlsi2601\\_1](https://doi.org/10.1207/s15327973rlsi2601_1)
- Kevallik, Leelo. 2018. What does embodied interaction tell us about grammar? *Research on Language and Social Interaction* 51(1). 1–21. DOI: <https://doi.org/10.1080/08351813.2018.1413887>
- Kendon, Adam. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26. 22–63. DOI: [https://doi.org/10.1016/0001-6918\(67\)90005-4](https://doi.org/10.1016/0001-6918(67)90005-4)
- Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511807572>
- Kendrick, Kobin H. & Holler, Judith. 2017. Gaze direction signals response preference in conversation. *Research on Language and Social Interaction* 50(1). 12–32. DOI: <https://doi.org/10.1080/08351813.2017.1262120>
- Kendrick, Kobin H. & Holler, Judith & Levinson, Stephen. 2023. Turn-taking in human face-to-face interaction is multimodal: Gaze direction and manual gestures aid the coordination of turn transitions. *Philosophical Transactions of the Royal Society B: Biological Sciences* 378(1875). 20210473. DOI: <https://doi.org/10.1098/rstb.2021.0473>
- Kolde, Raivo. 2019. pheatmap: Pretty heatmaps. R-package. <https://cran.r-project.org/web/packages/pheatmap/index.html>
- Konrad, Reiner & Hanke, Thomas & Langer, Gabriele & Blanck, Dolly & Bleicken, Julian & Hofmann, Ilona & Jeziorski, Olga & König, Lutz & König, Susanne & Nishio, Rie & Regen, Anja & Salden, Uta & Wagner, Sven & Worsack, Satu & Schulder, Marc. 2020. MY DGS – annotated. Public Corpus of German Sign Language, 3rd release. DOI: <https://doi.org/10.25592/dgs.corpus-3.0>
- Koole, Tom & Gosen, Myrte N. 2024. Scopes of reciprocity: An organization of responses to informings. *Journal of Pragmatics* 222. 25–39. DOI: <https://doi.org/10.1016/j.pragma.2024.01.004>
- Landis, J. Richard & Koch, Gary G. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159. DOI: <https://doi.org/10.2307/2529310>

- Lepeut, Alysso & Shaw, Emily. 2022. Time is ripe to make interactional moves: Bringing evidence from four languages across modalities. *Frontiers in Communication* 7. 780124. DOI: <https://doi.org/10.3389/fcomm.2022.780124>
- Levinson, Stephen. 2015. Other-initiated repair in Yélf Dnye: Seeing eye-to-eye in the language of Rossel Island. *Open Linguistics* 1(1). 386–410. DOI: <https://doi.org/10.1515/opli-2015-0009>.
- Lindblad, Gustaf & Allwood, Jens. 2015. Multimodal communicative feedback in Swedish. In *Proceedings of the 2nd European and the 5th Nordic symposium on multimodal communication*, 53–59. <https://ep.liu.se/ecp/110/008/ecp15110008.pdf>.
- Loos, Cornelia & Steinbach, Markus & Repp, Sophie. 2024. Polar response strategies across modalities: Evidence from German Sign Language (DGS). *Language* 100(3). 433–467. DOI: <https://doi.org/10.1353/lan.2024.a937185>
- Lutzenberger, Hannah & Wael, Lierin De & Omardeen, Rehana & Dingemans, Mark. 2024. Interactional infrastructure across modalities: A comparison of repair initiators and continuers in British Sign Language and British English. *Sign Language Studies* 24(3). 548–581. DOI: <https://doi.org/10.1353/sls.2024.a928056>
- Malisz, Zofia & Włodarczak, Marcin & Buschmeier, Hendrik & Skubisz, Joanna & Kopp, Stefan & Wagner, Petra. 2016. The ALICO corpus: Analysing the active listener. *Language Resources and Evaluation* 50, 411–442. DOI: <https://doi.org/10.1007/s10579-016-9355-6>
- Manrique, Elizabeth. 2016. Other-initiated repair in Argentine Sign Language. *Open Linguistics* 2(1). 1–34. DOI: <https://doi.org/10.1515/opli-2016-0001>
- Manrique, Elizabeth & Enfield, Nick. 2015. Suspending the next turn as a form of repair initiation: Evidence from Argentine Sign Language. *Frontiers in Psychology* 6. 1326. DOI: <https://doi.org/10.3389/fpsyg.2015.01326>
- Marmorstein, Michal & Szczepek Reed, Beatrice. 2023. Newsmarks as an interactional resource for indexing remarkability: A qualitative analysis of Arabic wallāhi and English really. *Contrastive Pragmatics* 5(1–2). 238–273. DOI: <https://doi.org/10.1163/26660393-bja10091>
- Maynard, Senko K. 1990. Conversation management in contrast: Listener response in Japanese and American English. *Journal of Pragmatics* 14(3). 397–412. DOI: [https://doi.org/10.1016/0378-2166\(90\)90097-W](https://doi.org/10.1016/0378-2166(90)90097-W)
- McCarthy, Michael. 2003. Talking back: “Small” interactional response tokens in everyday conversation. *Research on Language & Social Interaction* 36(1). 33–63. DOI: [https://doi.org/10.1207/S15327973RLSI3601\\_3](https://doi.org/10.1207/S15327973RLSI3601_3)
- Mesch, Johanna. 2016. Manual backchannel responses in signers’ conversations in Swedish Sign Language. *Language & Communication* 50. 22–41. DOI: <https://doi.org/10.1016/j.langcom.2016.08.011>
- Mol, Lisette & Krahmer, Emiel & Maes, Alfons & Swerts, Marc. 2011. Seeing and being seen: The effects on gesture production. *Journal of Computer-Mediated Communication* 17(1). 77–100. DOI: <https://doi.org/10.1111/j.1083-6101.2011.01558.x>
- Mondada, Lorenza. 2016. Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics* 20(3). 336–366. DOI: [https://doi.org/10.1111/josl.1\\_12177](https://doi.org/10.1111/josl.1_12177)

- Mori, Taiga & Jokinen, Kristiina & Den, Yasuharu. 2022. Cognitive states and types of nods. In Paggio, Patrizia & Gatt, Albert & Tanti, Marc (eds.), *Proceedings of the 2nd workshop on people in vision, language, and the mind*, 17–25. Marseille, France: European Language Resources Association. <https://aclanthology.org/2022.pvlam-1.4/>.
- Navarretta, Costanza & Paggio, Patrizia. 2010. Classification of feedback expressions in multimodal data. In *Annual meeting of the association for computational linguistics, 48th ACL*, 318–324. Uppsala, Sweden: Association for Computational Linguistics.
- Navarretta, Costanza & Paggio, Patrizia. 2012. Multimodal behaviour and feedback in different types of interaction. In *Proceedings of the eighth international conference on language resources and evaluation (LREC)*, 2338–2342. Istanbul, Turkey: European Language Resources Association (ELRA).
- Omardeen, Rehana. 2023. *Providence Island Sign Language in interaction*: Georg-August-University Göttingen PhD Thesis. DOI: <https://doi.org/10.53846/goediss-10243>
- Oomen, Marloes & Roelofsen, Floris. 2023. Biased polar question forms in Sign Language of the Netherlands (NGT). *FEAST. Formal and Experimental Advances in Sign language Theory* 5, 156–168. DOI: <https://doi.org/10.31009/FEAST.i5.13>
- Özyürek, Aslı. 2021. Considering the nature of multimodal language from a crosslinguistic perspective. *Journal of Cognition* 4(1). 42. DOI: <https://doi.org/10.5334/joc.165>
- Perniss, Pamela. 2018. Why we should study multimodal language. *Frontiers in Psychology* 9. 1109. DOI: <https://doi.org/10.3389/fpsyg.2018.01109>
- Puupponen, Anna. 2019. Towards understanding nonmanuality: A semiotic treatment of signers' head movements. *Glossa* 4(1). 39. DOI: <https://doi.org/10.5334/gjgl.709>
- R Core Team. 2025. R: A language and environment for statistical computing. <https://www.R-project.org/>.
- Rasenberg, Marlou & Pouw, Wim & Özyürek, Aslı & Dingemanse, Mark. 2022. The multimodal nature of communicative efficiency in social interaction. *Scientific Reports* 12(1). 19111. DOI: <https://doi.org/10.1038/s41598-022-22883-w>
- Rasenberg, Marlou & Özyürek, Aslı & Dingemanse, Mark. 2020. Alignment in multimodal interaction: An integrative framework. *Cognitive Science* 44. e12911. DOI: <https://doi.org/10.1111/cogs.12911>
- Rossano, Federico & Brown, Penelope & Levinson, Stephen. 2009. Gaze, questioning, and culture. In Sidnell, Jack (ed.), *Conversation Analysis: Comparative perspectives* (Studies in Interactional Sociolinguistics), 187–249. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511635670.008>
- Sacks, Harvey & Schegloff, Emanuel A. & Jefferson, Gail. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4). 696–735. DOI: <https://doi.org/10.2307/412243>
- Safar, Josefina & De Vos, Connie. 2022. Pragmatic competence without a language model: Other-initiated repair in Balinese homesign. *Journal of Pragmatics* 202. 105–125. DOI: <https://doi.org/10.1016/j.pragma.2022.10.017>

- Sandler, Wendy. 2024. Speech and sign: The whole human language. *Theoretical Linguistics* 50(1–2). 107–124. DOI: <https://doi.org/10.1515/tl-2024-2008>
- Sbranna, Simona & Möking, Eduardo & Wehrle, Simon & Grice, Martine. 2022. Backchannelling across languages: Rate, lexical choice and intonation in L1 Italian, L1 German and L2 German. In *Proc. speech prosody*, 734–738. DOI: <https://doi.org/10.21437/SpeechProsody.2022-149>
- Schegloff, Emanuel A. 1968. Sequencing in conversational openings. *American Anthropologist* 70(6). 1075–1095. DOI: <https://doi.org/10.1525/aa.1968.70.6.02a00030>
- Schegloff, Emanuel A. 1982. Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In Tannen, Deborah (ed.), *Analyzing discourse: Text and talk*, 71–93. Washington, D.C.: Georgetown University Press.
- Schloerke, Barret & Cook, Di & Larmarange, Joseph & Briatte, Francois & Marbach, Moritz & Thoen, Edwin & Elberg, Amos & Toomet, Ott & Crowley, Jason & Hofmann, Heike & Wickham, Hadley. 2024. GGally: Extension to 'ggplot2'. <https://cran.r-project.org/web/packages/GGally/index.html>.
- Schulder, Marc & Hanke, Thomas. 2022. How to be FAIR when you CARE: The DGS Corpus as a case study of open science resources for minority languages. In Calzolari, Nicoletta & Béchet, Frédéric & Blache, Philippe & Choukri, Khalid & Cieri, Christopher & Declerck, Thierry & Goggi, Sara & Isahara, Hitoshi & Maegaard, Bente & Mariani, Joseph & Mazo, Hélène & Odiijk, Jan & Piperidis, Stelios (eds.), *Proceedings of the thirteenth language resources and evaluation conference*, 164–173. Marseille, France: European Language Resources Association (ELRA).
- Selting, Margret & Auer, Peter & Barth-Weingarten, Dagmar & Bergmann, Jörg & Bergmann, Pia & Birkner, Karin & Couper-Kuhlen, Elizabeth & Deppermann, Arnulf & Gilles, Peter & Günthner, Susanne & Hartung, Martin & Kern, Friederike & Mertzluft, Christine & Meyer, Christian & Morek, Miriam & Oberzaucher, Frank & Peters, Jörg & Quasthoff, Uta & Schütte, Wilfried & Stukenbrock, Anja & Uhmman, Susanne. 2009. Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10. 353–402.
- Simon, Carsta. 2018. The functions of active listening responses. *Behavioural Processes* 157. 47–53. DOI: <https://doi.org/10.1016/j.beproc.2018.08.013>
- Skedsmo, Kristian. 2020. Other-initiations of repair in Norwegian Sign Language. *Social Interaction: Video-Based Studies of Human Sociality* 3(2). DOI: <https://doi.org/10.7146/si.v3i2.117723>
- Skedsmo, Kristian. 2023. Repair receipts in Norwegian Sign Language multiperson conversation. *Journal of Pragmatics* 215. 189–212. DOI: <https://doi.org/10.1016/j.pragma.2023.07.015>
- Stivers, Tanya. 2008. Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language & Social Interaction* 41(1). 31–57. DOI: <https://doi.org/10.1080/08351810701691123>
- Stubbe, Maria. 1998. Are you listening? Cultural influences on the use of supportive verbal feedback in conversation. *Journal of Pragmatics* 29(3). 257–289. DOI: [https://doi.org/10.1016/S0378-2166\(97\)00042-8](https://doi.org/10.1016/S0378-2166(97)00042-8)
- Tolins, Jackson & Fox Tree, Jean E. 2014. Addressee backchannels steer narrative development. *Journal of Pragmatics* 70. 152–164. DOI: <https://doi.org/10.1016/j.pragma.2014.06.006>

- Tottie, Gunnel. 1991. Conversation style in British and American English: The case of backchannels. In Aijmer, Karin & Altenberg, Bengt (eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*, 254–271. London: Longman.
- Trujillo, James P. & Simanova, Irina & Bekkering, Harold & Özyürek, Asli. 2018. Communicative intent modulates production and comprehension of actions and gestures: A Kinect study. *Cognition* 180. 38–51. DOI: <https://doi.org/10.1016/j.cognition.2018.04.003>
- Trujillo, James P. & Vaitonyte, Julija & Simanova, Irina & Özyürek, Asli. 2019. Toward the markerless and automatic analysis of kinematic features: A toolkit for gesture and movement research. *Behavior Research Methods* 51(2). 769–777. DOI: <https://doi.org/10.3758/s13428-018-1086-8>
- Truong, Khiet P. & Poppe, Ronald & Kok, Iwan De & Heylen, Dirk. 2011. A multimodal analysis of vocal and visual backchannels in spontaneous dialogs. In *Proc. Interspeech 2011*, 2973–2976. DOI: <https://doi.org/10.21437/Interspeech.2011-744>
- Uhmann, Susanne. 1996. On rhythm in everyday German conversation: Beat clashes in assessment utterances. In Couper-Kuhlen, Elizabeth & Selting, Margret (eds.), *Prosody in conversation*, 303–365. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511597862.010>
- van Gijn, Rik & Hirtzel, Vincent & Gipper, Sonja & Ballivián Torrico, Jeremías. 2011. The Yurakaré Archive. <https://hdl.handle.net/1839/8df587ed-3d6e-4db8-bfe5-4ecad5cef3a2>.
- Vandenitte, Sébastien. 2023. When referents are seen and heard: A comparative study of constructed action in the discourse of LSF (French Belgian Sign Language) signers and Belgian French speakers. In Gardelle, Laure & Vincent-Durroux, Laurence & Vinckel-Roisin, Hélène (eds.), *Reference: From conventions to pragmatics*, 127–149. Amsterdam/Philadelphia: John Benjamins. DOI: <https://doi.org/10.1075/slcs.228.07van>
- Vigliocco, Gabriella & Perniss, Pamela & Vinson, David. 2014. Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1651). 20130292. DOI: <https://doi.org/10.1098/rstb.2013.0292>
- White, Sheida. 1989. Backchannels across cultures: A study of Americans and Japanese. *Language in Society* 18(1). 59–76. <http://www.jstor.org/stable/4168001>. DOI: <https://doi.org/10.1017/S0047404500013270>
- Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis. 2nd edition*. Cham: Springer. DOI: <https://doi.org/10.1007/978-3-319-24277-4>
- Wiener, Morton & Devoe, Shannon. 1974. Regulators, channels, and communication disruption. Clark University unpublished research proposal.
- Wiltschko, Martina. 2021. *The grammar of interactional language*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/9781108693707>
- Xu, Jun. 2016. *Displaying reciprocity: Reactive tokens in Mandarin task-oriented interaction*. Amsterdam/Philadelphia: John Benjamins. DOI: <https://doi.org/10.1075/scld.6>

Yngve, Victor H. 1970. On getting a word in edgewise. In *Chicago Linguistics Society, 6th Meeting (CLS-70)*, 567–577. Chicago, Illinois, USA: University of Chicago.

Zellers, Margaret. 2021. An overview of forms, functions, and configurations of backchannels in Ruruuli/Lunyala. *Journal of Pragmatics* 175. 38–52. DOI: <https://doi.org/10.1016/j.pragma.2021.01.012>

