HANNAH LUTZENBERGER,
LIERIN DE WAEL,
REHANA OMARDEEN,
AND MARK DINGEMANSE

Interactional Infrastructure across Modalities: A Comparison of Repair Initiators and Continuers in British Sign Language and British English

Abstract

Minimal expressions are at the heart of interaction: Interjections like "Huh?" and "Mhm" keep conversations flowing by establishing and reinforcing intersubjectivity among interlocutors. Crosslinguistic

Hannah Lutzenberger's research interests combine language variation, language emergence, and language acquisition. She obtained her PhD from Radboud University in Nijmegen, the Netherlands, for her work on Kata Kolok phonology and is currently a postdoc at Tilburg University, the Netherlands. Lierin de Wael is a speech therapist, audiologist, and linguist with a strong interest in specific language impairment and hearing impairment in young children. She is currently based in the Netherlands. Rehana Omardeen obtained her PhD from the University of Göttingen through work documenting and describing aspects of Providence Island Sign Language and currently works as a policy officer at the European Union of the Deaf. Mark Dingemanse studies how language is shaped by and for social interaction. He is based at the Centre for Language Studies, Radboud University in Nijmegen, the Netherlands.

research has identified that similar interactional pressures can yield structurally similar words (e.g., to initiate repair across languages). While crosslinguistic comparisons that include signed languages remain uncommon, recent work has revealed similarities in discourse management strategies among signers and speakers that share much of their cultural background. This study contributes a crossmodal comparison of repair initiators and continuers in speakers of English and signers of British Sign Language (BSL). We combine qualitative and quantitative analyses of data from sixteen English speakers and sixteen BSL signers, resulting in the following: First, the interactional infrastructure drawn upon by speakers and signers overwhelmingly relies on behaviors of the head, face, and body; these are used alone or sometimes in combination with verbal elements (i.e., spoken words or manual signs), while verbal strategies alone are rare. Second, discourse management strategies are remarkably similar in form across the two languages: A held eye gaze or *freeze-look* is the predominant repair initiator and head nodding the main continuer. These results suggest a modality-agnostic preference for visual strategies that do not occupy the primary articulators, one that we propose is founded in recipiency; people maintain the flow of communication following principles of minimal effort and minimal interruption.

Introduction

Minimal expressions are at the heart of conversation: Interjections like "Huh?" and "Mhm" keep conversations flowing by establishing and reinforcing intersubjectivity among interactants (Dingemanse 2024). These expressions have gained recent prominence in the study of spoken languages. For example, crosslinguistic research has shown that similar interactional environments can yield structurally similar words across unrelated languages (Dingemanse, Torreira, and Enfield 2013). In research on signed languages, too, there is growing interest in interactional phenomena such as backchanneling (e.g., Fenlon, Schembri, and Sutton-Spence 2013; Mesch 2016) and repair initiation (e.g., Byun et al. 2017; Manrique 2016; Skedsmo 2020a; Safar and de Vos 2022; Omardeen 2023).

While there has been increasing descriptive work on individual signed languages, crosslinguistic comparisons that include signed languages on par with spoken languages remain uncommon (Okrent 2002; Taub, Galvan, and Piñar 2009, Enfield et al. 2013, Floyd et al.

2016). Nevertheless, a handful of recent studies have examined discourse management strategies among signers and speakers from the same culture, revealing striking similarities (Lepeut and Shaw 2022). For example, French Belgian Sign Language (LSFB) signers and Belgian French speakers use manual holds in similar ways to manage turn-taking (Lepeut 2020). Studies like this highlight the fundamental role of crossmodal comparisons in understanding universalities in the interactional infrastructure and unraveling shared and modality-specific elements of discourse management.

In this study, we contribute one such crossmodal comparison using data from speakers of British English and signers of British Sign Language (BSL). We investigate minimal expressions (or interjections) in two environments that are fundamental to communication: continuers and repair initiators. *Continuers* are displays of understanding that invite the other to go on, encouraging progressivity in conversation, while *repair initiators* work to flag trouble, halting progressivity until some trouble is resolved. While this interactional work can be accomplished in many different ways, here, we zero in on the most minimal forms that these interactional resources take in each of the languages. Using conversational data from both languages, we explore similarities and differences in form in two well-defined sequential contexts.

The article is structured as follows: We first provide an overview of the two interactional resources studied in this article, continuers and repair initiators, and discuss relevant literature, focusing on minimal expressions. We then turn to the current study, explaining the methodology, including data, data annotation, and analyses, followed by presenting the results in terms of descriptive statistics and qualitative discussion of the form of continuers and repair initiators. We end the article with a discussion of our results as they relate to aspects of universality and efficiency and discuss limitations and future work.

Continuers and Repair Initiators

Continuers

Continuers are highly frequent words or short utterances that display an understanding that some unit has been received and that more is to come (Goodwin 1986). There is considerable diversity in terminology when it comes to continuers; they have also been called

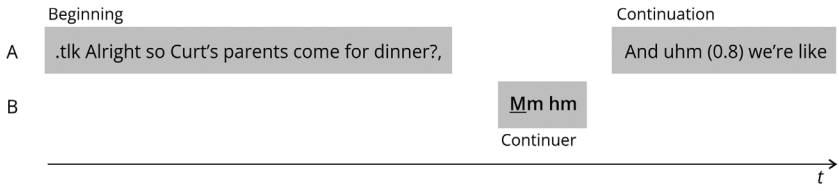

FIGURE 1. A common sequential context for continuers, illustrated using an excerpt adapted from Stivers (2008, 41).

backchannel responses (Yngve 1970), *listener responses* (McCarthy et al. 2002), *response tokens* (Gardner 2001), *minimal responses* (Fellegy 1995) and *reactive tokens* (Young and Lee 2004) among other labels. For the purpose of this study, we refer to them as continuers (Goodwin 1986). Despite the terminological variety, researchers agree that engaged recipiency is a key element of everyday social interaction, and continuers are a key way in which recipients participate and collaborate in conversation.

Continuers typically provide feedback without requesting the conversational floor (O'Keeffe and Adolphs 2008). This can be observed in the sequential positioning of continuers, as illustrated in figure 1, where participant A begins a telling by producing a turn-constructional unit with an upwardly intoned final element (represented in the transcript by "?,"), projecting there is more to come. After a brief silence, story recipient B offers a vocal continuer "mm hm," aligning with the telling activity, after which A provides a continuation (Stivers 2008). Continuers often appear at points where a turn has reached possible completion yet where more by the same speaker is projected. Some such points coincide with places where floor transition could be relevant; others fall within the ongoing turn (Kjellmer 2009; Lambertz 2011; Howes and Eshghi 2021).

In spoken languages, continuers can take on a range of forms, from minimal vocalizations like "Mhm" to fuller lexical and phrasal forms like "Yeah," "Good," or "That's right," the latter often associated with doing additional interactional work such as mirroring stance (O'Keeffe and Adolphs 2008). In this article, we focus on the most minimal form made available in the respective languages and do not investigate phrasal continuers. A recent crosslinguistic study provides

the following modality-agnostic description of the form of continuers: "Optimal continuers are (1) easy to plan and produce, (2) unobtrusive, and (3) sufficiently distinct from regular words to be seen as ceding the conversational floor" (Dingemanse, Liesenfeld, and Woensdregt 2022). That work finds great similarity in the form of these expressions across a sample of thirty-two spoken languages. Reports of visual elements associated with the continuer function in spoken languages are common; specifically, blinks, head nodding, smiling, and shifting eye gaze have been reported as continuers in some spoken languages (e.g., Kendon 1967; Stivers 2008; Knight 2009; Knutson 2009; Hömke, Holler, and Levinson 2017).

Continuers have also been studied in some signed languages. Mesch (2016) analyzed continuers (backchannels) in approximately thirty-five minutes of conversational data of sixteen signers of Swedish Sign Language (SSL). This study showed that most continuers (80 percent) are realized nonmanually through "nodding, head-shaking, smiling, change of body posture, nose wrinkles and widened eyes" (Mesch 2016, 40), with only 20 percent involving manual signs. Similarly, a study on BSL investigated the effect of sociolinguistic variation on continuers, coding manual and nonmanual continuers, with the latter category only head nods (Fenlon, Schembri, and Sutton-Spence 2013). They found that age affects the use of manual continuers, a finding also corroborated by the SSL study: Younger signers produce fewer manual continuers than older signers (Mesch 2016; Fenlon, Schembri, and Sutton-Spence 2013). While an early BSL study by Sutton Spence (2000; cf. Fenlon, Schembri, and Sutton-Spence 2013) suggested an effect of gender on backchanneling (male BSL signers backchannel less than female BSL signers), this was not supported in Fenlon and colleagues' (2013) more recent study. Lastly, Mesch (2016) identified in SSL another possibly modality-specific form of continuers, termed *weak manual activity* (see figure 2). Weak manual activity refers to small movements of the hand or fingers of a signer, in the low or even outside of the signing space, often produced in the lap; this activity is not generally classified as lexical sign due to the divergence from the language-specific phonological standards, and Mesch (2016, 41) proposes parallels to interjections such as "Uh-huh" in spoken languages.

FIGURE 2. Example of weak manual activity as a continuer is produced with minimal effort in the signer's lap in SSL. Image from Mesch (2016, 35). Reproduced with permission from Johanna Mesch.

Repair Initiators

Other-initiated repair refers to the system of interactional resources people use to address trouble in producing, perceiving, or interpreting conversational exchanges. They typically feature a "question-like action" (Floyd et al. 2016, 178) that singles out a prior turn as troublesome and in need of repair. Repair sequences halt productivity, requiring both parties to cooperate to signal the trouble and provide a resolution before the conversation can go forward. Prototypically, repair sequences are composed of the following structure (figure 3): A repair-initiating turn by participant B (here, "Huh?") retrospectively

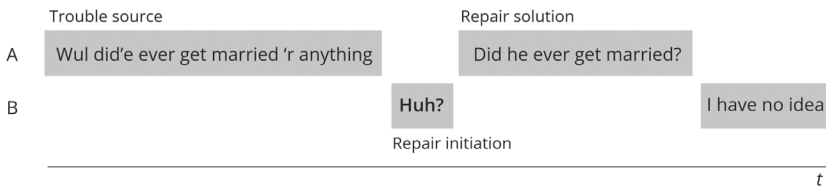

FIGURE 3. The sequential structure of other-initiated repair, illustrated using an excerpt adapted from Schegloff, Jefferson, and Sacks (1977, 367).

designates a prior turn by the other as troublesome, and invites a repair solution in next position. Often, a single repetition or elaboration is sufficient to restore the conversation, as here, though sometimes reaching closure requires several attempts (Skedsmo 2020b). Here, we are concerned only with the repair-initiating turn.

Repair initiation generally comes in three basic formats: restricted request, restricted offer, and open request (Dingemanse and Enfield 2015). Restricted repair formats use question words or partial repetition to single out a particular element of the prior turn as troublesome and lie beyond the scope of this article. Open repair formats do not specify the problem source and target the entire previous turn. They come in a range of formats, from interjections ("Huh?") to formulaic expressions like "Sorry?" or fuller questions like "What did you say?" (Clift 2016). For the purposes of this article, we focus on minimal expressions of the open repair type only, excluding any phrasal or formulaic expressions.

Minimal expressions in open repair sequences are mostly interjections like "Huh?" and have been found to be strikingly similar across unrelated spoken languages, possibly due to convergent cultural evolution (Dingemanse, Torreira, and Enfield 2013). Further, some studies have shown the relevance of manual gestures in repair sequences (e.g., Rasenberg et al. 2022; Mortensen 2016; Sikveland, Ove, and Ogden 2012; Jokipohja and Lilja 2022), and others mention co-occurring facial and bodily behavior (e.g., Floyd et al. 2016; Goodwin 2000; Seo and Koshik 2010), such as eyebrow activity, shifts in eye gaze, and head and body movements (Clift 2016; Enfield et al. 2013). Nevertheless, few studies to date have investigated verbal and nonverbal behaviors in conjunction, and evidence for nonverbal elements remains limited to isolated reports such as in examples from Yélî Dnye, a Papuan language used on Rossel Island, where recipients may initiate repair solely through "an intense gaze and a freezing of response" (Levinson 2015, 404 and see Manrique and Enfield 2015).

In recent years, the body of work on other-initiated repair in signed languages has been growing. The literature includes studies on Argentine Sign Language (LSA; Manrique 2016), Norwegian Sign Language (NTS; Skedsmo 2020a), Providence Island Sign Language (PISL; Omardeen 2023), cross-signing (situations in which two deaf

people who do not share a language meet; Byun et al. 2017), and Balinese homesign (Safar and de Vos 2022). In these languages, signers have been documented using several means of open repair initiation, including question signs, mouthing of question words, and various nonmanual behaviors. These nonmanuals include movements of the eyebrows (e.g., raising, lowering), the torso (e.g., leaning forward or backward), the head (e.g., tilting forward, backward), the mouth (e.g., pouching lips), and the eyes (e.g., squinting). Moreover, in several languages, signers have been noted to use *freeze-looks* to initiate repair; a freezing of the face and body while maintaining eye contact with the active signer (Manrique and Enfield 2015). Overall, nonmanual behaviors in signed languages have been compared to spoken interjections, operating on the same level of "minimal effort and questioning prosody" (Manrique 2016, 5–6).

This Study

Continuers and repair initiators are fundamental to communicative interactions. It appears that all languages have interactional resources devoted to these functions, and moreover that the resources they recruit show similarities in form and function (Dingemanse 2024). With the growing availability of multimodal corpora of everyday interactions in spoken and signed languages, there is now an opportunity to address two challenges: (1) exploring the role of visual cues comprehensively, including manual gestures and other visual behavior, and (2) carrying out a crossmodal comparison by including data from both signed and spoken languages.

The present study contributes a crossmodal comparison of continuers and repair initiators in two populations who share many aspects of their cultural background: hearing speakers of English (most of whom are British) and deaf signers of BSL. Specifically, we ask, "What forms do repair initiators and continuers take in each language?" and "Do forms overlap?" In asking these questions, we set out to investigate universality in interactional infrastructure by drawing on evidence from both speakers and signers. Focusing on nonphrasal continuers and repair initiators only, we combine qualitative and quantitative analyses of sixty-four minutes of signed and spoken conversations, cross-sampling eight minutes from dyadic conversations of each of the

sixteen speakers of (British) English from the Rossi Corpus of Conversational English (Rossi and Kendrick 2013) plus two newly collected dyads and sixteen BSL signers from the BSL Corpus (Schembri et al. 2017). We identified instances of continuers and repair initiators in both datasets based on function and analyzed their form. Our parallel study applies the same method and the same close coding with the same criteria to two different datasets. As such, it complements recent crossmodal comparisons (Lepeut 2020; Lepeut and Shaw 2022; Shaw 2019) by demonstrating similarities in the minimal features of discourse management in interaction across modalities.

Methods

In this study, we combine quantitative and qualitative data analysis of sixty-four minutes of signed and spoken conversations. Specifically, we focus on minimal formats of continuers and repair initiators (open requests), excluding any phrasal expressions and including all minimal behaviors articulated by the voice, hands, face, head, and body. In the following, we first describe the two datasets we base our study on and our coding criteria and then give an overview of the quantitative and qualitative analyses.

Data

The data used in this study stem from two different sources: the BSL Corpus (Schembri et al. 2017) and the Rossi Corpus of Conversational English (Rossi and Kendrick 2013) plus two newly collected dyadic conversations (see appendix A for details).

Signed dataset. The BSL Corpus contains data of 249 deaf BSL signers, recorded at eight different locations across the United Kingdom between 2008 and 2011 and includes a variety of different (semi-)spontaneous tasks (Schembri et al. 2017). For the purpose of this study, we randomly selected eight conversations from the BSL Corpus, focusing on dyadic, conversational data between two deaf signers. Each session lasts around thirty minutes.

Spoken dataset. The Rossi Corpus of Conversational English comprises fifteen video recordings with a total duration of 411 minutes (Rossi and Kendrick 2013). Participants are native speakers of English,

predominantly British English, who were videotaped in the United Kingdom in informal settings such as at the participants' homes, university facilities, cafés, etc. Each session lasts between seven-and-a-half to forty minutes and varies as to the number of interlocutors. For this study, we included only dyadic conversations ($N = 6$) to match the conversational setting available in the BSL Corpus. Given this criterion, we recorded two additional dyadic conversations among native speakers of British English in Amsterdam, the Netherlands.

From each dataset, we selected eight dyadic conversations. For practical reasons, segments of eight minutes per conversation were annotated and analyzed, totaling 64 minutes of data per language.

Data Coding

Data was coded using ELAN (Crasborn and Sloetjes 2008; ELAN 2020). BSL data was coded by HL and English data was coded by LdW. To ensure a valid comparison and avoid circularity, we identify forms in clearly defined sequential environments using the sequential control method (Floyd 2021). For continuers and repair initiators, we kept track of several structural dimensions and formal characteristics, coding the sequence, target, content, and the existence and form of the different relevant communicative behaviors in separate tiers.

For continuers, we coded all recipient behavior that was not immediately followed by a floor transition or attempted floor transition. For repair initiators, we identified sequences of other-initiated repair and included only open repair formats, those that request repair without specifying what or where the trouble is. For both categories, we included only minimal instances, excluding any phrasal expressions. For both continuers and repair initiations, we classified instances by their form into three categories: verbal only, visual only, or combined. For the purposes of this study, we use verbal to refer to vocalizations or standalone mouthings and manual articulations. Most commonly, this would be spoken words for English speakers and manual signs for BSL signers. We use visual to refer to all other visible communicative actions, including movements of the hands (for English speakers only), head, face, mouth, and body (i.e., all visual behavior that does not occupy the primary articulators). For English speakers, these behaviors

have been referred to as *co-speech gesture* and *nonverbal behavior* in prior literature. In BSL, these behaviors have been referred to as *nonmanual* behavior.

A note on the coding of mouthing by BSL signers: For the purpose of this study, we consider all mouthing (Boyes-Braem and Sutton-Spence 2001; Bank, Crasborn, and van Hout 2011; Crasborn et al. 2008) as part of the verbal category. Mouthing is often redundant, in which case the meaning of the mouthed word matches exactly the meaning of the manual sign; these cases were not coded. Nonredundant mouthing is also possible, where the mouthing does not match the manual sign but adds information; these were coded in the verbal category. Finally, mouthing can occur on its own with no accompanied manual sign; these are also included in the verbal category, as we feel they align both in form and meaning more with spoken words and manual signs than they do with visual behaviors. Nonredundant mouthings were always coded independently of other co-occurring nonmanual elements.

Coding proceeded in three rounds; initially, we identified all candidates of continuers and repair initiators. We then added the linguistic coding and double-checked all instances, and finally, HL and LdW reviewed all instances together and discussed unclear cases in the data. We used the export-function embedded in ELAN to extract data and conducted all analyses in R (R Core Team 2019).

Analysis

We investigate the research questions: What forms do repair initiators and continuers take in each language, and do forms overlap? Based on prior work showing great crosslinguistic similarity in the form of verbal continuers and repair initiators, we predict to find forms and expressions that resemble what has previously been described for different languages. Moreover, we expect some cross-language similarities across our two datasets, especially in visual forms, given the shared cultural background and interactional environments. To answer these questions and test these predictions, we conducted quantitative and qualitative analyses for each dataset, all of which are descriptive in nature. First, we analyzed the frequency of different patterns (visual only, verbal only, combined). Then, we investigated the data more

Table 1. Summary of descriptive results

	BSL	English
Continuers	520	450
Repair initiators	44	56

qualitatively, analyzing the exact forms within each of the attested categories. We conducted each of these analyses separately per language first and then highlighted similarities and differences.

Results

Overview Descriptive Results

The data yielded 970 instances of continuers and 100 instances of repair initiators across both datasets; we identified 520 instances of continuers in BSL and 450 in English, and forty-four instances of repair initiators in BSL and fifty-six in English (see table 1). In the following sections, we focus first on continuers and then on repair initiators, report results for both languages individually, and then draw a comparison. For both, we present an overview of the quantitative results, providing raw numbers and frequencies in percentages, and then we discuss the data more qualitatively by elaborating on examples. Although raw frequencies are hardly informative (Schegloff 1993), it may be useful to note that continuers appear to be about ten times as frequent as repair initiators, with about seven to eight occurrences per minute in the datasets considered here.

Continuers

Quantitative Results: Frequencies of Patterns. The following section provides an overview of quantitative results. The findings from our two datasets are summarized in table 2. Continuers could (but did not always) take the form of visual and/or verbal elements in both languages.

Continuers in BSL were most commonly only visual elements (68.5 percent; $N = 356/520$) and never occurred as only a manual sign. Indeed, combining visual and verbal cues was a less common strategy for continuers in BSL (31.5 percent; $N = 164/520$).

Table 2. Overview of continuers in BSL and English

	continuers						
	visual		verbal		visual + verbal		**total**
	n	%	n	%	n	%	
BSL	356	68.5%	0	0%	164	31.5%	**520**
English	198	44%	100	22%	152	33.8%	**450**

In English, all three forms were attested. Most frequently, English speakers produced visual continuers (44 percent; $N = 198/450$). Also common was the combination of visual and verbal elements for continuers (33.8 percent; $N = 152/450$) and continuers that were only a word (i.e., verbal-only continuers) were used less often (22 percent; $N = 100/450$).

Across both datasets, continuers almost always included visual elements. They were most commonly only visual in both BSL and English or they combined visual and verbal elements. Interestingly, continuers that took only a visual form were more frequent in BSL than in English, while combinations of visual and verbal elements were slightly more frequent in English than in BSL. Purely verbal forms were attested exclusively in English; continuers that are manual signs only were unattested in BSL.

Qualitative Results: Examples of Different Forms. In the following section, we present qualitative insights into all attested forms in BSL and English.

In our BSL dataset, we found two forms of continuers: purely visual ones and the ones combining visual and verbal elements (see figure 4). Purely visual continuers most often consisted of a single visual element (57 percent; $N = 203/356$) and less frequently, they combined multiple visual elements (42.9 percent; $N = 153/356$).

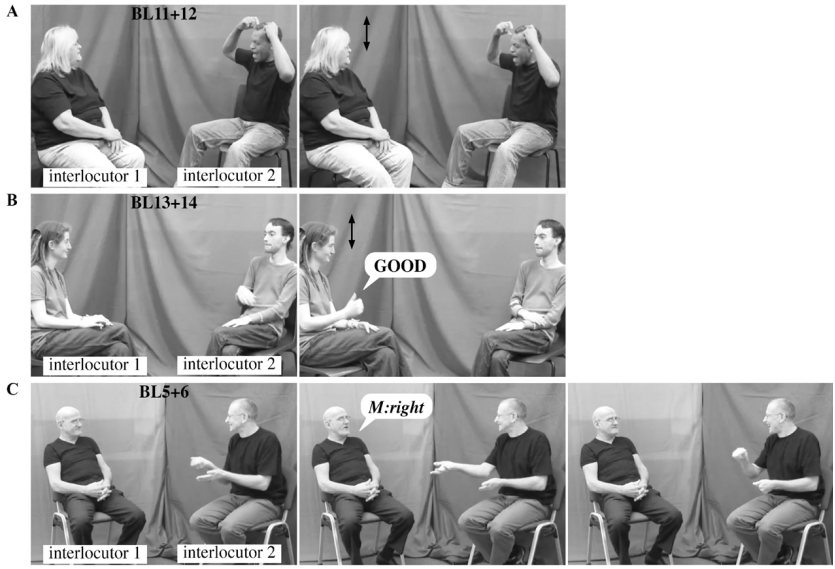

FIGURE 4. Examples of continuers in BSL: (A) nonverbal continuer head nod (Interlocutor 1); (B) combined continuer head nod and manual sign GOOD (Interlocutor 1); (C) nonverbal continuer nonredundant mouthing "Right" (Interlocutor 1).

Combined visual and verbal continuers, on the other hand, more often consisted of multiple visual elements (58.5 percent; $N = 96/164$) than a single visual element (41.5 percent; $N = 68/164$). By far, the most common form of continuer in BSL was a head nod (92.7 percent; $N = 482/520$), irrespective of whether the visual element cooccurred with a verbal element. Much less frequent than head nodding were raised eyebrows ($N = 72$) and smiling ($N = 59$), followed by an open mouth ($N = 36$), tilting the head backward ($N = 32$), shaking the head ($N = 24$), or tilting the head sideward ($N = 20$). All other nonmanual elements occurred with a frequency lower than twenty tokens and might therefore be considered somewhat less systematic. An interesting aspect of our results from BSL lies in the continuers combining visual and verbal elements. While we found that signers often use manual signs such as RIGHT ($N = 23$), GOOD ($N = 17$), and BAD ($N = 11$), we also found that signers often use the mouth for what we classified as verbal elements; in roughly half the cases in which signers combine visual and verbal elements for continuers,

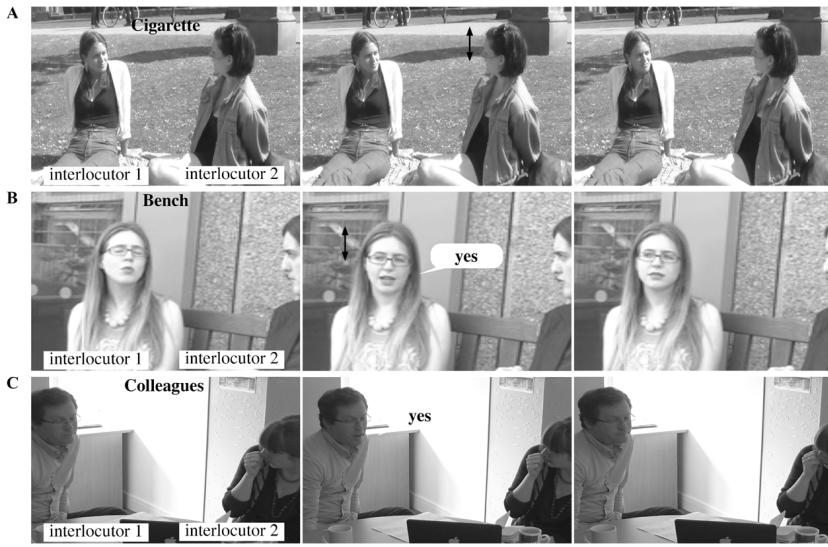

Figure 5. Examples of continuers in English: (A) visual-only continuer realized by head nod (Interlocutor 2); (B) combined continuer realized by a head nod and "Yes" (Interlocutor 1); (C) verbal-only continuer "Yes" (Interlocutor 1).

the signer produced a nonredundant mouthing (49.2 percent; N = 32/65). Specifically, these standalone mouthings took the form and functions of spoken continuers in our dataset, most commonly the mouthings "Ah" (N = 23), "Ya" (N = 17), "Right" (N = 15), "Yes" (N = 11), and "Yeah" (N = 10).

In our English dataset, continuers that used only visual elements, only verbal elements, or combined verbal and visual elements were attested (see figure 5). Out of the 198 instances of visual-only continuers, the vast majority (78.8 percent; N = 156/198) included a single visual element, and only a fifth of the data (20.7 percent; N = 41/198) featured multiple visual elements. Similarly, thirty-two of the instances of continuers that combined visual and verbal elements featured multiple visual elements (21.1 percent; N = 32/152), but most combined instances included only a single visual element (78.9 percent; N = 120/152). In both categories, head nods were the most frequent (N = 143). Much less frequent than head nodding were other head movements (N = 92), laughter (N = 49), and tilting the head backward (N = 24). All other visual elements occurred with a

frequency lower than twenty tokens and might, therefore, be considered somewhat less systematic. In terms of combined and purely verbal continuers, "Yeah" (N = 81) and "Mm" (N = 70) were the most prototypical forms.

In sum, we found striking similarities across both datasets. First, continuers with a visual component most frequently featured a single visual element rather than multiple ones. Second, this visual element was almost always a head nod. Third, besides a head nod, there were two visual elements that occurred frequently among English speakers and BSL signers, namely tilting the head backward and smiling in BSL and laughing in English. The latter observation is interesting given that the coding revealed that smiling is much less common among English speakers (N = 9) and laughter much less common among BSL signers (N = 6). We believe that this is caused by coding definitions, defining *laughter* as an expression of happiness with sound and smiling as a silent equivalent; it is interesting, however, that this seems to reflect the primary mode of articulation of English speakers (i.e., auditory-vocal input and output) and BSL signers (i.e., visual-spatial input and output), and could be investigated further in the future. Fourth, the verbal forms across English speakers and the nonredundant, standalone mouthings were not identical but very similar in that they took minimal forms.

Repair Initiators

Quantitative Results: Frequencies of Patterns. Repair initiators in our datasets most commonly included only visual elements; in some instances, BSL signers and English speakers combined visual and verbal elements, but instances without visual elements were virtually unattested. The findings from both datasets are summarized in table 3.

Repair initiators in BSL consisted predominantly of only visual elements; in 77.3 percent (N =3 4/44) of the data, only visual elements were used to initiate repair. In the remaining portion of the data, BSL signers combined visual elements with verbal ones (22.7 percent; N = 10/44). They never used verbal elements alone for this function.

In English, all three formal categories were attested, yet our data revealed a very strong preference for initiating repair with visual

Table 3. Overview of repair initiators in BSL and English

	repair initiator						
category	visual		verbal		visual + verbal		**total**
	n	%	n	%	n	%	
BSL	34	77.3%	0	0%	10	22.7%	**44**
English	50	89.3%	1	1.7%	5	8.9%	**56**

elements only (89.3 percent; $N = 50/56$). Involving verbal elements was considerably less common; while combining visual and verbal elements (8.9 percent; $N = 5/56$) was attested more frequently, using exclusively verbal elements for this purpose (1.8 percent; $N = 1/56$) was rare.

Across both datasets, English speakers and BSL signers showed a strong preference for visual-only forms of repair initiators. Indeed, this tendency was even stronger among the English speakers than among the BSL signers. In turn, only in the English dataset did we find one instance in which a participant produced a repair initiator exclusively verbally and without any visual elements.

Qualitative Patterns per Language: Examples of Different Forms. In the following section, we provide a qualitative discussion of the forms found across the data.

In the BSL data, all repair initiators were either purely visual or combined visual and verbal elements (see figure 6); no verbal-only repair invitations were attested. Most cases included a single nonmanual element; multiple nonmanual elements were coded only in 15 percent ($N = 15/100$) of the sample, three instances of combined visual and verbal repair initiators, and twelve instances of purely visual repair initiators. Across both categories, the most prototypical visual element used to initiate repair was clearly the freeze-look ($N = 29$) (Manrique 2016). Three of these include other nonmanuals: In one instance, the freeze-look co-occurred with shifting the eye gaze, and

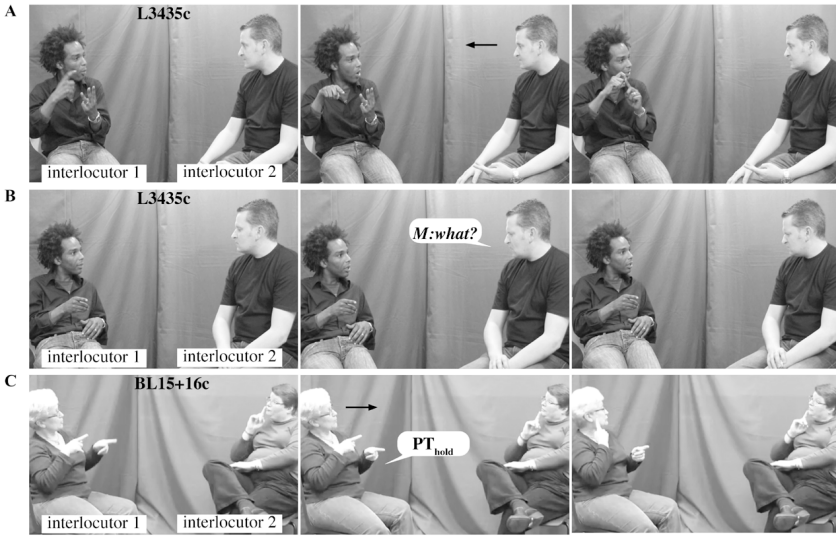

FIGURE 6. Examples of repairs in BSL: (A) visual-only repair initiator realized by a freeze-look (Interlocutor 2); (B) combined repair initiator realized standalone mouthing "what?" and furrowed eyebrows (Interlocutor 2); (C) combined repair initiator realized by held manual pointing sign alongside a freeze-look (Interlocutor 1).

in two other instances, the freeze-look was produced together with an open mouth. Besides freeze-looks, two additional nonmanual elements that occurred much less often were raised eyebrows ($N = 7$) and open mouth ($N = 5$). All other nonmanual elements that were attested occurred with a frequency lower than five tokens. The ten instances of combined repair initiators included a range of different signs; seven instances were cases of manual holds of a sign, two were instances of the lexical sign WHAT, and one instance was a standalone nonredundant mouthing (namely the mouthing "What?"). Interestingly, six instances of combined repair initiators involved freeze-looks; in all those cases, the verbal component of the repair initiator was a hold of a manual sign (e.g., AGE, TWO, pointing sign [PT], etc).

In the English dataset, most repair initiators were purely visual, but we also found a few combined instances and one purely verbal one (see figure 7). Most of the visual-only cases included multiple visual elements (69.6 percent; $N = 39/56$) rather than consisting of

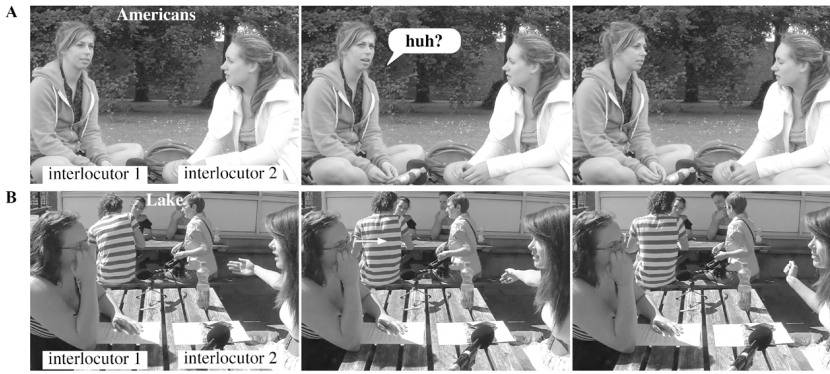

Figure 7. Examples of repairs in English: (A) combined repair initiator "Huh?" alongside furrowed eyebrows (Interlocutor 1); (B) visual-only repair initiator realized by a freeze-look (Interlocutor 1).

a single visual element (30.4 percent; $N = 17/56$). Although the overall number was low, when repair initiators combined verbal and visual elements, a single visual element was more frequent (60 percent; $N = 3/5$) than multiple ones (40 percent; $N = 2/5$). When it comes to the form of the visual elements, the pattern is clear: Freeze-looks were the most common form of repair initiators among English speakers ($N = 36$), followed by furrowed eyebrows ($N = 17$), and further, multiple blinks ($N = 6$) and an open mouth ($N = 6$). In this dataset, freeze-looks were more often accompanied by other visual elements ($N = 26/36$), particularly blinking, than they occurred by themselves ($N = 10/36$). All other attested visual elements occurred in less than five tokens and were therefore regarded as less systematic and not reported here. In terms of combined instances, the forms of the verbal component include only three types: "What?," "Huh?," and "Mhm," all of which are also found in prior crosslinguistic comparisons (Dingemanse and Enfield 2015). No systematicity could be found across the five instances of combined repair initiators in terms of what forms combined; from those five instances, it did not seem like particular verbal and visual cues usually go together. The only token of a solely verbal repair initiator involved the use of the general question word *what*.

Discussion and Conclusion

In this study, we conducted a crossmodal comparison of continuers and repair initiators in two populations that share a cultural background: speakers of (British) English and signers of BSL. Our study reveals two main results. First, the The interactional infrastructure drawn upon by both speakers and signers overwhelmingly relies on visual resources; repair initiations and continuers are rarely exclusively verbal (i.e., spoken words or manual signs) and less often combine verbal visual elements. Second, discourse management strategies are remarkably similar in form across the two languages: Head nodding is the main continuer format and a freeze-look is the predominant repair initiator format. These results suggest a modality-agnostic preference for visual strategies, one that we propose is founded in recipiency; speakers and signers maintain the flow of communication following the principles of minimal effort and minimal interruption.

Multimodality for Minimizing Joint Effort: Interactional Infrastructure Overwhelmingly Relies on Visual Resources

The interactional infrastructure drawn upon by English speakers and BSL signers overwhelmingly relies on visual resources, that is, visible movements of the face, head, and body that do not occupy the primary articulators. We have shown that for continuers and repair initiation, speakers and signers most commonly recruit visual elements alone or in combination with verbal elements; very rarely, signers and speakers use exclusively verbal behaviors, that is, spoken words or manual signs.

A major contribution of this direct, side-by-side comparison of form and frequency is the prevalence across modalities of visual forms. When it comes to signed languages, we know that nonmanual elements are integrated into linguistic structure at varying levels and are fundamental to signed language communication. However, this study emphasizes the crucial role and high frequency of visual elements in spoken language communication. In research on spoken languages, visual behaviors (also called *co-speech gesture nonverbal behavior*) have been studied in the realm of continuers and repair initiators, but by using parallel datasets and parallel coding and analyses, our results

reveal just how similar in frequency and in form these cues are across signed and spoken languages. We also found that when verbal forms were used, they were overwhelmingly accompanied by visual forms, with both signers and speakers exploiting multiple articulators at once to give feedback in conversation. This emphasizes the crucial role of visual elements that do not occupy the primary articulators in these functional niches.

Why are these visual articulations so central to the conversational infrastructure? One reason may be that they minimize interruptions to the speaker and offer a streamlined way due to their easy perceptibility. This may be particularly true when it comes to visual expressions of the face and head, where efficiency is permitted by joint eye gaze. In conversation, the active speaker monitors their interlocutor, periodically establishing mutual gaze for feedback during their turn (Kendon 1967; Goodwin 1984). In signed languages, mutual gaze is arguably even more important to feedback and, in fact, despite the hands being the "primary" articulators, signers have been shown to focus eye gaze on their interlocutor's face rather than their hands (e.g., Bosworth, Stone, and Hwang 2020). Given this attention to the face, minimal expressions of the face and head may be a highly efficient way to capitalize on that visual attention and signal misunderstanding (in the case of repair) or agreement/interest/encouragement to continue (in the case of continuers). Where verbal interjections may be perceived as an attempt to take the turn, visual signaling may be a more noninvasive feedback mechanism to allow the speaker to self-correct or continue their turn. Another reason may be that they are easy to produce quickly and with minimal effort. Nods and kin cost far less time and effort to produce than manual signs. This is particularly true when it comes to facial nonmanuals: The hands are relatively large articulators that may take more preparation to move with slower reaction time, whereas the mouth, eyebrows, eyes, and other facial articulators are all considerably smaller.

Taken together, these may also provide some explanation for the use of standalone nonredundant mouthings we have found in our dataset for both repairs and continuers (also reported in NTS repair sequences in Skedsmo 2020a). By reducing a unit that combines manual activity and redundant mouthing to only a mouthing, the

informational load shifts from both the hands and the mouth to only the mouth. This reduction alters the productive and perceptive load, minimizing the potential for interrupting the conversational flow. This makes nonredundant standalone mouthings great candidates for efficient continuers and repair initiators.

Our results are in line with research showing that multimodal cues are at the heart of communication and are key tools for displaying recipiency and calibrating understanding (Holler 2022). Continuers and repair initiators both serve to streamline the flow of conversation, and it makes sense that their forms would be adapted to this purpose. We have shown that in both BSL and spoken English, people frequently recruit visual and specifically nonmanual elements. These visual elements signal understanding or misunderstanding and halt or encourage progressivity in conversations. Perhaps there is a special benefit in the particular realization of these forms, in line with prior work showing that people can communicate more efficiently by exploiting the simultaneity of multimodal cues (Rasenberg et al. 2022). In this way, the forms we have observed may be optimally suited to aid communicative efficiency (Levshina and Moran 2021).

In sum, multimodality allows for simultaneity and overlap of communicative behaviors, by enabling people to perceive the other's facial expressions while simultaneously talking/signing. These routinized practices allow them to easily navigate simultaneity with minimal expressions like the ones attested in this study. Our results underscore the importance of multimodality in face-to-face conversation (Hamilton and Holler 2023) and the need for a thoroughly multimodal perspective on language (e.g., Vigliocco, Perniss, and Vinson 2014; Perniss 2018; Holler and Levinson 2019).

Universality: Discourse Management Strategies Are Remarkably Similar in Form across Both Languages

The preference for multimodal elements and for recruiting similar elements of English speakers and BSL signers suggests that universal pressures of face-to-face conversation push languages in different modalities towards using visible articulators in the same ways in discourse management. Furthermore, the strategies are remarkably similar in form across both languages: In both British English and BSL, head

nodding is the most frequent continuer, and freeze-looks are the most common repair initiator. Thus, we suggest that signers and speakers draw on the same resources for achieving the fundamentals of interaction. Regardless of language modality, signers and speakers rely on the same set of behaviors of the face, head, and body to manage conversation.

Why is it that these elements take the same form? Continuers and repair initiators both represent structural sequences and communicative functions that are found across different languages. This means that similar pressures are at play across different languages, possibly leading to similar forms that emerge and stabilize as a result of convergent cultural evolution. For example, a continuer like "Mhm" is frequent across different languages because the same pressures, such as ease of production and minimal interruption, exist across different languages. Indeed, Dingemanse and colleagues (2022, 166) argue that "a combination of effort minimization and a relaxation of the pressure for the reuse of phonetic features can push continuer-like words to the same low-effort yet distinctive part of the possibility space across languages." When it comes to repair, universal pressures may also contribute to the prevalence of freeze-looks across BSL and English. As Floyd and colleagues (Floyd 2016) demonstrate, holds can play a key part in repair sequences. They examined Northern Italian, Cha'palaa, and Argentine Sign Language, finding that across languages the timing of holds is remarkably similar with respect to the timing of the problem and solution parts of the repair sequence. Their conclusion, that similarities in visual bodily practices may also emerge from similar conversational pressures (Floyd et al. 2016), aligns with our findings of freeze-looks as the most prevalent repair strategy among BSL signers and English speakers.

While we would expect that similar expressions are found across different languages, given the shared ecological niche, we would also expect some crosslinguistic variation, for instance, due to cultural differences. Specifically, we predict not only that the distribution of expressions would be similar across language pairs, where speakers/signers share the same cultural background, but also that there is room for cultural variations between languages. Our data offers a reasonably matched sample, with most (but not all) English speakers being

from the United Kingdom and all BSL signers being from the United Kingdom: This largely shared overlap in culture may well account for the striking similarities in the forms that discourse management strategies take. Indeed, if a third language was added to the comparison from a culture where a head nod was used for negation and a head shake for affirmation, one would expect differences in the preferred continuer form for that language, be it spoken or signed.

Limitations and Future Work

This study complements recent direct crossmodal comparisons (Lepeut 2020; Lepeut and Shaw 2022; Shaw 2019) by demonstrating similarities in the fundamental features of discourse management in interaction across modalities. We demonstrated striking similarities in the form of continuers and repair initiators in English speakers and BSL signers and there is reason to expect further similarities across languages. First, prior work on other-initiated repair finds considerable overlap in the form of verbal and visual elements of repair initiators in unrelated languages (Enfield et al. 2013). Second, findings from Mesch (2016) on STS suggest considerable form overlap in continuers with the current dataset. Both pieces of evidence might further support the potential for universality discussed here. Having said this, this study represents only two points of comparison, and more studies of speakers and signers of different languages need to be undertaken to test how widely these findings can be generalized.

Our study has revealed strong parallels between the most frequent visual elements used as continuers and repair initiators. Our sample focused on English speakers and BSL signers where overlap in the forms may also be influenced by culture. Nevertheless, it needs to be stressed that our sample not only includes mostly speakers of British English but also two speakers of American English, and that recordings for the data from the Rossi Corpus of Conversational English took place in the United Kingdom, while the two newly collected conversations were recorded in the Netherlands. While it is possible that our results are somewhat influenced by this speaker diversity, we have opted for retaining these data, given the clear results from our study and that this is the first parallel comparison undertaken. Still, future studies could be more rigorous in using only data from speakers

and signers who share all aspects of their cultural background to rule out any influence of this factor on the results. Furthermore, large-scale crosslinguistic studies including (signed and spoken) languages from distinct gestural cultures can help us understand potential universal trends in the forms of visual behaviors used for discourse management, similar to what has been done for verbal behaviors in spoken languages (Floyd 2021).

We find that freeze-looks present a number of challenges: (1) difficulties with identifying and coding freeze-looks as a result of how they are defined, (2) the fact that freeze-looks represent a complex and varied set of visual elements, and (3) the suggestion that freeze-looks might be better understood as holds, as suggested by Floyd and colleagues (2016). First, freeze-looks are defined as follows: "When a person has just been asked a direct question, instead of answering the question in the next turn position, the person holds still while looking directly at the questioner" (Manrique and Enfield 2015, 1). This definition is thus founded on the absence of any expressions, and this may create challenges in identifying and coding freeze-looks. Specifically, it may be hard to identify a freeze-look as there are no clear guidelines for how long such a lack of reaction lasts, what should be considered a start or end point, or how it may or may not be resolved. Second, freeze-looks are treated as a visual element but are defined through a functional sequence in conversation. Unlike the name suggests, they are not simply a function of eye gaze and often encompass a complex and varied set of held visual elements. For example, in a comparative study of holds across one signed and two spoken languages, Floyd and colleagues (2016) note that both manual elements and visual elements, including head, gaze, and posture, are all held meaningfully in repair sequences. These reasons may make it difficult to compare freeze-looks to the use of other visual elements, which are generally discussed as individual behaviors and described by their form (e.g., backward lean, smile, squint, all of which may be held during a freeze-look). Third, a useful alternative to the term *freeze-looks* can be found in Floyd and colleagues (2016), where these behaviors are referred to more generally as *holds*, that is, practices "in which relatively dynamic movements are temporarily and meaningfully held static" (Floyd et al. 2016, 176). Although this does

not address some of the issues mentioned above, it allows us a more holistic view of holds to include both manual and visual behaviors, instead of classifying visual holds as freeze-looks and manual holds as holds. This, in turn, allows us to explore the function and timing of holds more accurately. Nevertheless, despite these technical challenges in defining and coding freeze-looks, our results clearly support the fundamental importance of holds in repair initiation, as put forth by Floyd and colleagues (2016).

All the verbal and visual elements we analyzed are minimal in some way. For the purpose of this study, we operationalized minimal expressions as below the level of a phrase but we also included instances where speakers/signers recruited multiple visual elements. For spoken interjections, scholars have argued for their particular phonology to make them stand out; they are short and phonemically distinct; for instance, continuers are often vowelless and nasal (Gardner 1997). In sign language linguistics, while interjections have been likened to certain nonmanual expressions, there is little precise investigation into what the range of interjections in signed languages look like for several reasons: (1) Interjection-like expressions are not often coded in sign language corpora; (2) there is a focus on manual signs and nonmanual elements are often regarded as somewhat less important (Puupponen 2019); and (3) few studies directly test characteristics like the duration of signs as is done for spoken interjections (but see Börstell, Hörberg, and Östling 2016; Börstell, Crasborn, and Schembri 2019 for data on the duration of manual signs). Furthermore, whether nonmanual elements should be understood as minimal by default remains to be determined. One observation specific to signed languages are the standalone mouthings reported in this study and previously for NTS (Skedsmo 2020a). We treat these cases as reduced forms, originating in a combined expression of manual sign and congruent mouthing whereby the manual activity has been dropped and only the mouthing remained. Nevertheless, more research is needed on these forms, their relations to other items, and their interactional environments.

Concluding Remarks

We have conducted a direct comparison between minimal expressions used as continuers and repair initiators by speakers of English and

signers of BSL. We show that visual forms that do not occupy the primary articulators are the preference of both speakers and signers, and, on top of that, we found striking similarity in form across the two languages: The most common continuers are head nods, and the most common repair initiators are freeze-looks. Although this study is limited in its data and the languages of investigation, we see this study as the first step towards more controlled crossmodal comparisons, allowing us to further explore the pressures shaping similarities and differences across languages. Our study contributes evidence that the interactional infrastructure of languages is shared across modalities.

Our results suggest that across languages, minimal expressions, both verbal and visual, are present in similar forms, likely due to shared crosslinguistic pressures of face-to-face interaction. While our study is one of just a handful that have compared interaction in signed and spoken languages side by side, it shows the feasibility of cross-modal typology and hopefully inspires more such work in the future. Using sequentially anchored comparisons allows us to reexamine, reframe, and refine existing categories used across sign and spoken language research traditions, ultimately moving us towards a more holistic understanding of the interactional infrastructure of language and communication.

Acknowledgments

This article is based on a student project of HL and LdW during their research for their master's degree at University of Amsterdam and was supervised by Vadim Kimmelman.

The copyrighted BSL data in this article were collected for the British Sign Language Corpus Project (BSLCP) at University College London, funded by the Economic and Social Research Council UK (RES-620-28-6001), and supplied by the CAVA repository (https://bslcorpusproject.org/cava). The data are copyrighted. The English data are from the Rossi Corpus of Conversational English, used with permission.

HL is supported by a European Research Council Advanced Grant (ERC-ADG 885220 SignMorph) awarded to Adam Schembri.

MD is funded by a Dutch Research Council grant (Vidi 016.vidi.185.205 Elementary Particles of Conversation).

Author Contribution

Conceptualization: HL, LdW, MD; Methodology: HL, LdW, MD; Data Coding: HL, LdW; Analysis: HL, LdW; Visualization: HL, MD; Writing—original draft: HL, RO; Writing—review and editing: HL, RO, MD.

Conflict of Interest

The authors declare no competing interests.

References

Bank, R., O. A. Crasborn, and R. van Hout. 2011. Variation in Mouth Actions with Manual Signs in Sign Language of the Netherlands (NGT). *Sign Language & Linguistics* 14 (2): 248–70. https://doi.org/10.1075/sll.14.2.02ban.

Börstell, C., O. Crasborn, and A. Schembri. 2019. Frequency, Duration, and Signing Rate in Three Sign Language Corpora. Poster presented at the 13th Theoretical Issues in Sign Language Research (TISLR) Conference, University of Hamburg, Germany. 26-28 September 2019.

Börstell, C., T. Hörberg, and R. Östling. 2016. Distribution and Duration of Signs and Parts of Speech in Swedish Sign Language. *Sign Language & Linguistics* 19 (2): 143–96. https://doi.org/10.1075/sll.19.2.01bor.

Bosworth, R., A. Stone, and S. Hwang. 2020. Effects of Video Reversal on Gaze Patterns during Signed Narrative Comprehension. *The Journal of Deaf Studies and Deaf Education* 25 (3): 283–97. https://doi.org/10.1093/deafed/enaa007.

Boyes-Braem, P., and R. Sutton-Spence. 2001. *The Hands Are the Head of the Mouth: The Mouth as Articulator in Sign Languages*. Hamburg: Signum.

Byun, K., C. de Vos, A. Bradford, U. Zeshan, and S. C. Levinson. 2017. First Encounters: Repair Sequences in Cross-Signing. *Topics in Cognitive Science* 10 (2): 314–34. https://doi.org/10.1111/tops.12303.

Clift, R. 2016. *Conversation Analysis*. Cambridge, UK: Cambridge University Press.

Crasborn, O., E. van der Kooij, D. Waters, B. Woll, and J. Mesch. 2008. Frequency Distribution and Spreading Behavior of Different Types of Mouth Actions in Three Sign Languages. *Sign Language & Linguistics* 11 (1): 45–67. https://doi.org/10.1075/sll.11.1.04cra

Crasborn, O., and H. Sloetjes. 2008. Enhanced ELAN Functionality for Sign Language Corpora. *Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, 39–43. Luxembourg: European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/08022.html.

Dingemanse, M. 2024. Interjections at the Heart of Language. *Annual Review of Linguistics* 10:257–77. https://doi.org/10.1146/annurev-linguistics-031422-124743.

Dingemanse, M., and N. J. Enfield. 2015. Other-Initiated Repair across Languages: Towards a Typology of Conversational Structures. *Open Linguistics* 1 (1): 96–118. https://doi.org/10.2478/opli-2014-0007.

Dingemanse, M., A. Liesenfeld, and M. Woensdregt. 2022. Convergent Cultural Evolution of Continuers (Mmhm). In *The Evolution of Language: Proceedings of the Joint Conference on Language Evolution (JCoLE)*, 160–167. Nijmegen, the Netherlands: Joint Conference on Language Evolution. https://doi.org/10.17617/2.3398549.

Dingemanse, M., F. Torreira, and N. J. Enfield. 2013. Is "Huh?" A Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items. *PLoS ONE* 8 (11): e78273. https://doi.org/10.1371/journal.pone.0078273.

ELAN [Computer Software]. 2020. Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics, The Language Archive. https://archive.mpi.nl/tla/elan.

Enfield, N. J., M. Dingemanse, J. Baranova, J. Blythe, P. Brown, T. Dirksmeyer, P. Drew, et al. 2013. Huh? What?–A First Survey in Twenty-One Languages. In *Conversational Repair and Human Understanding*, ed. M. Hayashi, G. Raymond, and J. Sidnell, 343–80. Cambridge, UK: Cambridge University Press.

Fellegy, A. 1995. Patterns and Functions of Minimal Response. *American Speech* 70:186–99.

Fenlon, J., A. Schembri, and R. Sutton-Spence. 2013. Turn-Taking and Backchannel Behaviour in BSL Conversations. Poster presented at the 11th Theoretical Issues in Sign Language Research (TISLR) Conference, University College, London, 10–13 July 2013.

Floyd, S. 2021. Conversation and Culture. *Annual Review of Anthropology* 50 (1): null. https://doi.org/10.1146/annurev-anthro-101819-110158.

Floyd, S., E. Manrique, G. Rossi, and F. Torreira. 2016. Timing of Visual Bodily Behavior in Repair Sequences: Evidence from Three Languages. *Discourse Processes* 53 (3): 175–204. https://doi.org/10.1080/0163853X.2014.992680.

Gardner, R. 1997. The Conversation Object Mm: A Weak and Variable Acknowledging Token. *Research on Language & Social Interaction* 30 (2): 131–56. https://doi.org/10.1207/s15327973rlsi3002_2.

Gardner, R. 2001. *When Listeners Talk: Response Tokens and Listener Stance*. Amsterdam: Benjamins.

Goodwin, C. 1984. Notes on Story Structure and the Organization of Participation. In *Structures of Social Action: Studies in Conversation Analysis*, ed. J. Maxwell Atkinson and J. Heritage, 225–46. Cambridge, UK: Cambridge University Press.

Goodwin, C. 1986. Between and Within: Alternative Sequential Treatments of Continuers and Assessments. *Human Studies* 9 (2/3): 205–17.

Goodwin, C. 2000. Action and Embodiment within Situated Human Interaction. *Journal of Pragmatics* 32 (10): 1489–522. https://doi.org/10.1016/S0378-2166(99)00096-X.

Hamilton, A. F. D. C., and J. Holler. 2023. Face2face: Advancing the Science of Social Interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 378 (1875): 20210470. https://doi.org/10.1098/rstb.2021.0470.

Holler, J. 2022. Visual Bodily Signals as Core Devices for Coordinating Minds in Interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 377 (1859):20210094. https://doi.org/10.1098/rstb.2021.0094.

Holler, J., and S. C. Levinson. 2019. Multimodal Language Processing in Human Communication. *Trends in Cognitive Sciences* 23 (8): 639–52. https://doi.org/10.1016/j.tics.2019.05.006.

Hömke, P., J. Holler, and S. C. Levinson. 2017. Eye Blinking as Addressee Feedback in Face-To-Face Conversation. *Research on Language and Social Interaction* 50 (1): 54–70. https://doi.org/10.1080/08351813.2017.1262143.

Howes, C., and A. Eshghi. 2021. Feedback Relevance Spaces: Interactional Constraints on Processing Contexts in Dynamic Syntax. *Journal of Logic, Language and Information* 30 (June): 331–62. https://doi.org/10.1007/s10849-020-09328-1.

Jokipohja, A-K., and N. Lilja. 2022. Depictive Hand Gestures as Candidate Understandings. *Research on Language and Social Interaction* 55 (2): 123–45. https://doi.org/10.1080/08351813.2022.2067425.

Kendon, A. 1967. Some Functions of Gaze-Direction in Social Interaction. *Acta Psychologica* 26: 22–63. https://doi.org/10.1016/0001-6918(67)90005-4.

Kjellmer, G. 2009. Where Do We Backchannel?: On the Use of Mm, Mhm, Uh Huh and Such Like. *International Journal of Corpus Linguistics* 14 (1): 81–112. https://doi.org/10.1075/ijcl.14.1.05kje.

Knight, D. 2009. A Multi-Modal Corpus Approach to the Analysis of Backchanneling Behaviour. PhD thesis, University of Nottingham. http://eprints.nottingham.ac.uk/10786/.

Knutson, E. M. 2009. On Being Heard: A Study of Listening Behavior in French Conversation. *The French Review* 82 (6): 1180–93t.

Lambertz, K. 2011. Back-Channelling: The Use of Yeah and Mm to Portray Engaged Listenership. *Griffiths Working Papers in Pragmatics and Intercultural Communication* 4 (1/2): 11–18.

Lepeut, A. 2020. Framing Language through Gesture - Palm-Up, Index Finger-Extended Gestures, and Holds in Spoken and Signed Interactions in French-Speaking and Signing Belgium." PhD thesis, Namur, Belgium: University of Namur.

Lepeut, A., and E. Shaw. 2022. Time Is Ripe to Make Interactional Moves: Bringing Evidence from Four Languages across Modalities. *Frontiers in Communication* 7 (April). https://doi.org/10.3389/fcomm.2022.780124.

Levinson, S C. 2015. Other-Initiated Repair in Yélî Dnye: Seeing Eye-to-Eye in the Language of Rossel Island. *Open Linguistics* 1 (1): 386–410. https://doi.org/10.1515/opli-2015-0009.

Levshina, N., and S. Moran. 2021. Efficiency in Human Languages: Corpus Evidence for Universal Principles. *Linguistics Vanguard* 7 (s3): 20200081. https://doi.org/10.1515/lingvan-2020-0081.

Manrique, E. 2016. Other-Initiated Repair in Argentine Sign Language. *Open Linguistics* 2 (1): 1–34. https://doi.org/10.1515/opli-2016-0001.

Manrique, E., and N. J. Enfield. 2015. Suspending the Next Turn as a Form of Repair Initiation: Evidence from Argentine Sign Language. *Frontiers in Psychology* 6 (September): 1–21. https://doi.org/10.3389/fpsyg.2015.01326.

McCarthy, M., R Reppen, S.M. Fitzmaurice, and D Biber. 2002. Good Listenership Made Plain. British and American Non-Minimal Response Tokens in Everyday Conversation. In *Using Corpora to Explore Linguistic Variation*, ed. R. Reppen, S. M. Fitzmaurice, and D. Biber, 49–71. Amsterdam: John Benjamins.

Mesch, J. 2016. Manual Backchannel Responses in Signers' Conversations in Swedish Sign Language. *Language & Communication* 50 (September): 22–41. https://doi.org/10.1016/j.langcom.2016.08.011.

Mortensen, K. 2016. The Body as a Resource for Other-Initiation of Repair: Cupping the Hand Behind the Ear. *Research on Language and Social Interaction* 49 (January): 34–57. https://doi.org/10.1080/08351813.2016.1126450.

O'Keeffe, A., and S. Adolphs. 2008. Response Tokens in British and Irish Discourse: Corpus, Context and Variational Pragmatics. In *Pragmatics & Beyond New Series*, ed. K. P. Schneider and A. Barron, 178: 69–98. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/pbns.178.05ok.

Okrent, A. 2002. A Modality-Free Notion of Gesture and How It Can Help Us with the Morpheme vs. Gesture Question in Sign Language Linguistics (Or at Least Give Us Some Criteria to Work with). In *Modality and Structure in Signed and Spoken Languages*, ed. R. P. Meier, K. Cormier, and D. Quinto-Pozos, 175–98. Cambridge, UK: Cambridge University Press.

Omardeen, R. 2023. Providence Island Sign Language in Interaction. PhD diss., Georg-August-Universität Göttingen. https://doi.org/10.53846/goediss-10243.

Perniss, P. 2018. Why We Should Study Multimodal Language. *Frontiers in Psychology* 9. https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01109.

Puupponen, A. 2019. Towards Understanding Nonmanuality: A Semiotic Treatment of Signers' Head Movements. *Glossa: A Journal of General Linguistics* 4 (1): 39. https://doi.org/10.5334/gjgl.709.

R Core Team. 2019. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rasenberg, M., W. Pouw, A. Özyürek, and M. Dingemanse. 2022. The Multimodal Nature of Communicative Efficiency in Social Interaction. *Scientific Reports* 12 (1): 19111. https://doi.org/10.1038/s41598-022-22883-w.

Rossi, G., and K. H. Kendrick. 2013. *The Rossi Corpus of Conversational English*. Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics.

Safar, J., and C. de Vos. 2022. Pragmatic Competence without a Language Model: Other-Initiated Repair in Balinese Homesign. *Journal of Pragmatics* 202 (December): 105–25. https://doi.org/10.1016/j.pragma.2022.10.017.

Schegloff, E. A. 1993. Reflections on Quantification in the Study of Conversation. *Research on Language & Social Interaction* 26 (1): 99–128. https://doi.org/10.1207/s15327973rlsi2601_5.

Schegloff, E. A., G. Jefferson, and H. Sacks. 1977. The Preference for Self-Correction in the Organization of Repair in Conversation. *Language* 53 (2): 361–82.

Schembri, A., J. Fenlon, R. Rentelis, and K. Cormier. 2017. *British Sign Language Corpus Project: A Corpus of Digital Video Data and Annotations of British Sign Language 2008-2014*. London: University College London. http://www.bslcorpusproject.org.

Seo, M-S., and I. Koshik. 2010. A Conversation Analytic Study of Gestures That Engender Repair in ESL Conversational Tutoring. *Journal of Pragmatics* 42 (8): 2219–39. https://doi.org/10.1016/j.pragma.2010.01.021.

Shaw, E. 2019. *Gesture in Multiparty Interaction*. Washington, DC: Gallaudet University Press.

Sikveland, R. O., and R. A. Ogden. 2012. Holding Gestures across Turns: Moments to Generate Shared Understanding. *Gesture* 12 (2): 166–99. https://doi.org/10.1075/gest.12.2.03sik.

Skedsmo, K. 2020a. Other-Initiations of Repair in Norwegian Sign Language. *Social Interaction. Video-Based Studies of Human Sociality* 3 (2). https://doi.org/10.7146/si.v3i2.117723.

Skedsmo, K. 2020b. Multiple Other-Initiations of Repair in Norwegian Sign Language. *Open Linguistics* 6 (1): 532–66. https://doi.org/10.1515/opli-2020-0030.

Stivers, T. 2008. Stance, Alignment, and Affiliation during Storytelling: When Nodding Is a Token of Affiliation. *Research on Language and Social Interaction* 41 (1): 31–57.

Taub, S. F., D. Galvan, and P. Piñar. 2009. The Role of Gesture in Cross-modal Typological Studies. *Cognitive Linguistics* 20 (1): 71–92. https://doi.org/10.1515/COGL.2009.004.

Vigliocco, G., P. Perniss, and D. Vinson. 2014. Language as a Multimodal Phenomenon: Implications for Language Learning, Processing and Evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 20130292. https://doi.org/10.1098/rstb.2013.0292.

Yngve, V. 1970. On Getting a Word in Edgewise. *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, University of Chicago, 567–77.

Young, R. F., and J. Lee. 2004. Identifying Units in Interaction: Reactive Tokens in Korean and English Conversations." *Journal of Sociolinguistics* 8:380–407.

Appendix A: Details of the Dataset.

Language	Video file	Gender Interlocutor1	Gender Interlocutor2	Dataset
BSL	BL05+06c	male	male	BSL Corpus
	BL07+08c	male	female	BSL Corpus
	BL09+BL10c	female	female	BSL Corpus
	BL11+BL12c	female	male	BSL Corpus
	BL12+BL14c	male	female	BSL Corpus
	BL15+BL16c	female	female	BSL Corpus
	L34+35c	male	male	BSL Corpus
	M13+14c	female	female	BSL Corpus
English	Cigarette	female	female	Rossi Corpus
	Two friends	female	male	Rossi Corpus
	Americans	female	female	Rossi Corpus
	Colleagues	female	female	Rossi Corpus
	Bench	female	female	Rossi Corpus
	Lake	female	female	Rossi Corpus
	Colleagues	male	male	Rossi Corpus
	CREA	female	male	newly collected
	PCH	female	female	newly collected