Discourse Processes

Timing of Visual Bodily Behavior in Repair Sequences: Evidence From Three Languages

Simeon Floyd, Elizabeth Manrique, Giovanni Rossi & Francisco Torreira

Accepted author version posted online: 22 Dec 2014.
Published online: 29 Apr 2015.

Submit your article to this journal ⤤

Article views: 568

View related articles ⤤

View Crossmark data ⤤

Citing articles: 17 View citing articles ⤤

Routledge
Taylor & Francis Group

Timing of Visual Bodily Behavior in Repair Sequences: Evidence From Three Languages

Simeon Floyd, Elizabeth Manrique, Giovanni Rossi, and
Francisco Torreira

Language and Cognition Department
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

This article expands the study of other-initiated repair in conversation—when one party signals a problem with producing or perceiving another's turn at talk—into the domain of visual bodily behavior. It presents one primary cross-linguistic finding about the timing of visual bodily behavior in repair sequences: if the party who initiates repair accompanies their turn with a "hold"—when relatively dynamic movements are temporarily and meaningfully held static—this position will not be disengaged until the problem is resolved and the sequence closed. We base this finding on qualitative and quantitative analysis of corpora of conversational interaction from three unrelated languages representing two different modalities: Northern Italian, the Cha'palaa language of Ecuador, and Argentine Sign Language. The cross-linguistic similarities uncovered by this comparison suggest that visual bodily practices have been semiotized for similar interactive functions across different languages and modalities due to common pressures in face-to-face interaction.

INTRODUCTION

Repair and Visual Bodily Behavior

The domain of "repair" covers a range of practices that people use to deal with problems of "speaking, hearing and understanding" in interaction (Schegloff,

---

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hdsp.

Jefferson, & Sacks, 1977, p. 361). The spoken turns of repair sequences have been studied extensively since the first work on talk in interaction (Schegloff et al., 1977) until the present (e.g., Hayashi, Raymond, & Sidnell, 2013). However, less attention has been devoted to the study of the visual bodily practices that occur simultaneously with the spoken turns in repair sequences. In this article we expand the study of other-initiation of repair into the visual domain by comparing data from video corpora of everyday conversation in three unrelated languages: Northern Italian, the indigenous language Cha'palaa of Ecuador, and Argentine Sign Language (*Lengua de Señas Argentina* [LSA]). Including a sign language helps to broaden the discussion of repair from problems of "speaking and hearing" to those of "signing and seeing" (Manrique, 2011) or, in modality-neutral terms, "producing and perceiving." This is a broad field of inquiry, but here we narrow our focus to one main new finding from our research regarding a practice called a "hold," in which relatively dynamic movements are temporarily and meaningfully held static: in repair sequences, the duration of holds is timed according to the duration of the sequence.

Holds can include a wide range of behaviors using the head, eyes, face, hands, or torso or any combination of these that may be held in a stationary position and meaningfully disengaged at particular moments in interaction (for a more detailed discussion, see Holds, below). Here we look at holds that accompany repair initiation, that is, when a speaker uses formats like "huh?" or "what?" or "who?" or their cross-linguistic equivalents to signal a problem. Our qualitative and quantitative analyses of the timing of holds show that when the speaker who initiates repair holds a static bodily position during and after his or her repair initiation, that speaker maintains the hold until sequence closure, that is, until the problem is solved. The hold acts as a display of the speaker's orientation to the ongoing unresolved status of the repair sequence, and the subsequent disengagement from the hold displays the problem has been solved and the progressivity of the interaction can resume. After some further background on repair and multimodal interaction in this section, we next present quantitative results based on timing measurements of a sample of cases from the three languages, and then analyze several cases in excerpts from the corpora in detail, before concluding in the final section with some connections to broader issues.

Multimodality in Conversational Interaction

Conversation analysts have paid some attention to nonverbal behavior in interaction (e.g., Goodwin, 1981, 2000; Schegloff, 1984, 1998) and are increasingly looking at the role of bodily actions with respect to spoken turns and sequences, for example, how pointing or the rearrangement of objects may project upcoming turns or transitions between activities (Mondada, 2006, 2007). There is some debate about how visual bodily behavior (or nonlinguistic behavior,

in sign languages) relates to turns-at-talk. Nonlinguistic elements are technically not part of the turn-taking system in that they are not necessarily governed by a "one at a time" rule, as spoken (or signed) turns are (cf. Sacks, Schegloff, & Jefferson, 1974; Lerner & Raymond, unpublished data). At the same time, nonverbal conduct feeds into sequence organization (Clark, 1996; Levinson, 2013, among others) and is contingent on its development in many ways. Our approach to this problem is to reserve "turn" for elements that are subject to the one-speaker-at-a-time constraint as formulated in Sacks et al. (1974) and to use the term "move" (Enfield, 2009; Goffman, 1981) for a broader type of sequential practices that include, but are not limited to, the elements in turns, along with a range of visual bodily practices. The variety of visual bodily behavior involved in conversational interaction is daunting, but zooming in on one specific sequence type, other-initiated repair (OIR) sequences,[1] helps make questions in this area more focused. This "natural control" method (Dingemanse & Floyd, 2014; Enfield et al., 2013; Stivers et al., 2009) keeps the sequential context constant to examine how semiotic practices are used similarly or differently for specific conversational actions by speakers of different languages—in this case, for the management and resolution of OIR sequences.

The bodily behaviors we cover under the definition of a "hold" are certainly not limited to OIR sequences. They appear to be associated with managing the unresolved status of sequences and their resolution more generally, particularly in question–answer sequences (Rossano, 2012; Rossano, Brown, & Levinson, 2009; Stivers & Rossano, 2010). Emerging work suggests similar patterns to those we observe in several languages besides those we compare here. Li (2014) describes cases in Mandarin Chinese interaction in which the questioner leans toward the addressee while asking a question and maintains the position until the response. Rossano (2012) uses Italian interaction data to demonstrate the relevance of gaze maintenance to sequence continuation and of gaze withdrawal to sequence closure. Sikveland and Ogden's (2012) work on Norwegian shows a relationship between the maintenance of "co-speech" gestures until sequence closure, which they describe in terms of achieving "shared understanding." Clark (2005) discusses similar practices related to the manipulation of objects in the context of practical joint activities, showing for example that when assembling TV stands or Lego models, English-speaking participants hold objects stationary in particular positions to indicate there is still some business to attend to. With respect to sign languages, Manrique (2011) describes the maintenance of LSA signs in relation to ongoing unresolved OIR

---

[1]Repair initiation can be divided into two types: self-initiated, in which speakers redo or reformulate all or part of a turn without any on-record prompting by their addressee, and other-initiated, in which an addressee signals a problem, typically in the next sequential position after the problem turn (Schegloff et al. 1977, pp. 363–364). Here we are interested in OIR.

sequences and their resolutions. Other researchers have made similar observations about other sign languages (Groeber & Pochon-Berger, 2014).

These studies all pick up on practices of holding a bodily configuration stationary and then disengaging it around a specific sequential position, each focusing on a single type of practice (eye gaze, manual gesture, etc.) in a single language. Here we go beyond both language-specific accounts and analyses restricted to specific visual bodily behaviors to propose a set of generalizations that apply to a range of semiotic resources in informal face-to-face interaction cross-linguistically. Our empirical findings are limited to the function of holds in repair sequences, but they lead to potentially broader implications, which we outline in the conclusion.

Structure of Repair Sequences

Both in terms of turns and moves, OIR sequences are organized by the principle of "conditional relevance" (Schegloff, 1968; Schegloff & Sacks, 1973) or the way in which the production of a "first pair part" by one party (e.g., a question) projects and normatively expects the subsequent production of an appropriate "second pair part" by another (e.g., an answer), together forming an "adjacency pair." Repair initiation is a question-like action that expects the provision of a solution in next position. Such an action suspends the progressivity of the ongoing sequence, displacing the appropriate next turn with an OIR sequence (an "insert sequence" when the trouble source is a first pair part; Schegloff, 2007) that must be completed before the main sequence can be resumed or closed. This yields the following structure, which we treat as definitional of an OIR sequence:

Participant A:	trouble source	T-1
Participant B:	repair initiator	T0
Participant A:	repair solution	T+1
Participant B:	(OIR sequence closure/	T+2
	resume sequence from T-1)	

In the schema above, a "repair initiator" turn, which we call "T0" (a "next turn repair initiator"; Drew, 1997; Schegloff, 1992; Schegloff et al., 1977), occurs after a trouble source turn, called "T − 1", and before a repair solution turn, called "T + 1", usually containing a reformulation or clarification of the trouble source in T − 1.[2] The motivation for calling the trouble source T − 1

and not beginning with number one is that trouble sources are only recognizable in restrospect in light of the repair initiation at T0 (Schegloff, 2007); what happens when there are more than one repair attempts in a sequence is explained below. The participant who initiates the repair sequence in T0 is referred to as "B", and the participant who produces the trouble source in T − 1 and the solution turn in T + 1 is referred to as "A". In many sequences there is also a turn after T + 1, which we call "T + 2", in which the sequence is closed, either through information uptake or simply by continuing the sequence left off at T − 1. The following excerpt from Cha'palaa shows an example of this basic structure; a question in line 1 (T − 1) is followed by repair initiation in 2 (T0), a repetition in 3 (T + 1), and an answer resuming progressivity in 4 (T + 2):

(1) CHSF2011_01_11S3_3018944

1 A   nan   vale-n   washu-che-e
       how.much be.worth-Q washu-tree-FOC
       how much are washu trees?   ←T–1

2 B   aa   ((turns head and gaze towards A, off camera))
       huh?   ←T0

3 A   washu-che-e   nan   vale-n
       washu-tree-FOC   how.much be.worth-Q
       washu trees, how much are they?   ←T+1

4 B   Wilsun junu, kaspele, ma dolar-tene ju-de-e-we   ((turns away from A))
       Wilson   there   before   one   dollar-only   be-PL-become-N.EGO
       at Wilson's place they give only one dollar   ←T+2

IMAGE 1A   Speaker A (off camera) asks "How much are washu trees?" Speaker B (on bench, in the middle) begins to turn his head and gaze toward speaker A as he says "huh?" He then holds this position while speaker A repeats "washu trees, how much are they?"

IMAGE 1B   Speaker A (off camera) provides a repair solution by repeating "washu trees, how much are they?" As speaker B finishes answering this question, saying "At Wilson's place they are giving just one dollar," he turns his head and gaze back to their original position.

Excerpt (1) also provides a good illustration of a hold. Speaker B turns his head toward speaker A when producing the repair initiator in T0 and then turns away when resuming progressivity after T + 1. This is the general structure of all cases included in the sample we consider in this article, with the qualification that in some cases speaker B may treat the first repair solution (or lack thereof) as unsuccessful, and so may engage in "pursuit" (Bolden, Mandelbaum, & Wilkinson, 2012; Pomerantz, 1984) with one or more T0 repair initiations (see Holds in Pursuit Sequences, below). In these cases, the first T + 1 effectively takes the status of T − 1 for the continuing sequence and the hold duration orients not just to the next turn but to the next turn that successfully resolves the repair sequence. This is expected under our analysis; if the hold is about unfinished business, we predict its disengagement to coincide with sequence closure and to span multiple initiations of repair in such pile-ups (Enfield, Drew, & Baranova, unpublished data). For this reason, we identified individual cases on the basis of closure, not number of repair initiations, allowing us to test whether hold duration increased as sequence duration increased. In the cases we collected in the three video corpora each of the turns was labeled as T − 1, T0, T + 1, or T + 2, as appropriate, allowing us to make standard time measurements using the onsets and offsets of these turns as reference points.

Holds

Our interest in the phenomenon of holds in the context of repair sequences stems from initial observations by Manrique (2011) that contrasts between static and dynamic positions used in unresolved repair sequences and their resolution by signers of LSA. Groeber and Pochon-Berger (2014) make similar observations regarding Swiss German Sign Language (see also Baker, 1977 on American Sign Language). These practices are particularly salient in sign languages, but we suspected that spoken language interaction might feature similar phenomena.

Initial examination of the spoken language corpora revealed comparable "holding" practices in the same sequential position such as that seen in Excerpt (1).

Consider the following excerpt from LSA in comparison with Excerpt (1) from Cha'palaa, above. Similarly to the Cha'palaa excerpt, here Signer B changes his head position at T0, in this case in a raised and tilted-back position, holding it there. This is accompanied by holding the eyebrows together combined with a wrinkling of the nose to mark the interrogative construction (somewhat comparably to the interrogative prosody observed in some spoken languages). The hold also includes a manual sign articulated with the right hand, the interrogative pronoun "what," which is held stationary together with the nonmanual signs and head position until after the repair solution.

(2) **CAS_553706[3]**

| 1 | A | PRO-3 DISCUSS WHAT EVERTHING OK | |
| | | **They discuss what {to do}, everything is OK,** | |

		OTHERS ARGUE G:a-lot [PRO1 NO G: nothing	T−1
		{but} others argue a lot, {then} I don't {want to participate}	
		h-up--------**H** (hold)	
		Q-ET--------------**H**	
		Q-NW---------------**H**	
2	B	PRO3 WHAT PRO3 WHAT----**H**	T0
		{Who} are the others?	

| 3 | A | PRO3-TWO SN-C PRO3-TWO | T+1 |
| | | **They both {from} (sign name of an organization)** | |

| 4 | B | AH PRO3-TWO | T+2 |
| | | **Ah, they both** | |

IMAGE 2A   "....*WHAT*...?," participant B, right, initiates repair in line 4 by using manual (...*what?*) and nonmanual signs (wrinkled nose, eyebrows together, and upward head tilt.) holding them until participant A solves the sequence.

---

[3]The conventions used here for LSA follow Johnston's (2010) annotation system for AUSLAN as well as conversation analytic transcription conventions adapted for sign language interaction. See the Appendix for further explanation of transcription and glossing annotations for all three languages.

IMAGE 2B   *'...BOTH...',* participant A, on the left, end of the solution turn. *"Ah, ...both",* participant B, on the right, displays information uptake and confirms the repair solution in line 5 by manual *(...both)* and nonmanual signs (head down and wide open mouth).

IMAGE 2C   Participant A, right, disengages hold position of nonmanuals first, maintaining manual hold position ("WHAT?").

IMAGE 2D   Participant A, right, disengages manual sign hold position to provide uptake (line 5). It disengages in second place, after disengaging nonmanuals markers (see image 2C).

As with most of the cases we considered, the hold in this particular excerpt does not have just one main salient element but rather consists of a complex of practices, each with its own attributes and affordances. With close analysis it is possible to observe that in Excerpt (2) the nonmanual markers are disengaged

before the manual sign, as the face moves into open-mouthed position for the display of information uptake in line 5 ('*Ah!*'). The manual sign is disengaged shortly afterward as B confirms the repair solution by repeating part of it in T + 2 ('*they both*'). As was mentioned above, formats that display information uptake make up the second main type of T + 2 turn with which OIR sequences are closed; the first type, consisting of formats that resume progressivity from T − 1, was illustrated in Excerpt (1). More generally, looking at Excerpts (1) and (2) side by side provides a good illustration of the similar patterns of hold timing across different modalities that motivated our comparison.

At this point we should further clarify exactly how we are applying the term "hold." The term has been applied with several different definitions, some of which would not cover the heterogeneous set of visual bodily practices like those seen in Excerpts (1) (gaze and head direction) and (2) (head position, manual and nonmanual sign). McNeill's (1992) annotation system for visual/gestural articulation includes hold as an optional phase that may occur during the progression from "preparation" through "stroke" and "retraction." This system has been applied to manual gesture by many scholars (e.g., Park-Doob, 2010 looking specifically at gestural holds). Linguists have also found this system useful for classifying the articulatory phases of sign languages (e.g., Kita, Gijn, & Hulst, 1998), the held material in this case being primarily linguistic signs.

Here we focus on holds of bodily configurations that come into position during a conversational turn and that are then held beyond the end of the turn. In terms of gestural phases, these correspond to the phase of "poststroke holds" (Kita, 1990; McNeill, 1992), in which the last part of a gesture or sign's stroke is held temporarily before retraction or further articulation. But the phenomenon of holding we are considering here is broader than that defined by McNeill as "any temporary cessation of movement without leaving the gesture hierarchy" (1992, p. 83). We find not only manual gestures (or manual signs, in the case of sign languages) to be meaningfully held in interaction but also other visual elements, like movements of the head, gaze, and posture.

In this article, a hold is defined as any meaningful maintenance of a stationary bodily configuration in contrast with a dynamic disengagement or retraction, regardless of the exact nature of the configuration. This broad use of the term "hold" includes manual poststroke holds and transitions between a held sign and a subsequent sign utterance, as well as behaviors involving gaze and head direction, facial expression and posture. Additionally, we consider cases where manual and nonmanual elements are held together like that seen in Excerpt (2). We argue that in repair sequences holding behavior of all these kinds is a significant way in which interactants display orientation to the not-yet-resolved status of the problem. The types of visual bodily behavior that speakers hold in such contexts are related to the maintenance and monitoring of joint attention in interaction, including moving the

eyes, head, or body in the direction of an interlocutor or signing or gesturing such that an interlocutor can observe. We further suggest that because these attentional behaviors have a "natural meaning" (cf. Enfield, 2009, p. 179; Enfield, Kita, & de Ruiter, 2007, p. 1735; Grice, 1957), this cross-linguistic association with mutual attention has allowed them to be semiotized similarly in different speech communities, such that comparable practices are reflected in data sets from three unrelated languages in two different modalities.

QUANTITATIVE RESULTS OF TIMING MEASUREMENTS

Sampling Procedure

To test our predictions about hold timing using quantitative methods, we identified a sample set of cases of repair sequences in each of our three corpora—defined by the sequential structure outlined in the previous section—to take timing measurements. The video corpora were collected by the authors in their respective field sites as part of the project Human Sociality and Systems of Language Use, directed by Nick Enfield. Each sample consisted of selections of 10 minutes each from 14 different recordings, totaling 140 minutes per language. The recordings are of everyday interactions from households or other informal contexts, including a mix of dyadic and multiparty interactions among family and friends of different ages and genders. Selections were taken varyingly from the beginnings, middles, and ends of recordings and were required to include relatively sustained conversation (there were no silences longer than a minute).

This sampling procedure was designed according to the "natural control" method that keeps sequence type—in this case OIR sequences—constant to compare communicative behavior across different languages. We located all OIR sequences in the sample by the structural properties described above, independently of any nonlinguistic behavior, and created annotations for them in the corpus using the ELAN transcription program (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). The languages differed somewhat in the frequency of repair sequences, with Italian and LSA at around 100 sequences in 140 minutes and Cha'plaaa at around 60. Within this set of OIR cases, LSA had the highest frequency of holds at 80/100, with Italian at 57/100 and Cha'palaa at 40/62. As discussed in more detail below in Sequences Without Holds, the higher rate of holds is related to the fact that joint visual attention is usually a prerequisite for holds in OIR sequences, and this is more constant in sign language interaction as compared to spoken languages.

Of the data set of OIR cases including holds, we randomly selected a subset of 40 cases from each language in which some bodily configuration (for the types included, see Types of Visual Bodily Behavior, below) was held by B past the

offset of T0, for a total of 120 cases.[4] Using ELAN, we created time-aligned annotations for the turns included in each case, recording the durations of the repair initiator (T0), the repair solution (T + 1), and its sequence-closing ratification (T + 2), if there was one. For the spoken languages, turn onset and offset were measured from the first vocal articulation,[5] whereas for LSA the measure was slightly different in that turn onset was measured from the beginning of the preparation phase, not from the beginning of the stroke. Although for many questions spoken turns are more comparable with the stroke phase, here it was important to measure from B's first evidence of A's forthcoming repair solution, which is already present during the preparation phase. This small coding difference, reflective of modality differences, means LSA holds are recorded as being held slightly longer into T + 1 than with spoken T + 1 cases, as generally signers waited to disengage holds until seeing some of the stroke phase; however, this difference does not change the nature of the general phenomenon.

In addition to the turns in the OIR sequences, we also created an annotation marking the duration of the visual bodily hold to compare its timing with the spoken or signed elements of the sequence. When different resources were used in combination, such as gazing, gesturing/signing, and leaning forward simultaneously, we measured the timing based on the disengagement of the longest-held element in the configuration.

Timing of Holds Relative to the Repair Sequence Closure

As noted above, OIR sequences may either close at T + 1—the repair solution— or at T + 2—in the next position. In this section, we examine the timing of the end of holds in each repair sequence to the start of T + 1 and to the start of T + 2 in cases in which the OIR sequence was closed at T + 2. Figure 1 shows density plots of timing of hold ends relative to the start of T + 1 (i.e., the repair solution) for each of the three languages examined. In all three languages, hold ends overwhelmingly occur within a time window of 0 to 4 s after the onset of T + 1, with the most frequent timings between approximately 700 ms and 1 s depending on the language. In a small number of cases, the hold was disengaged close to or even previously to the onset of T + 1, suggesting the timing of hold ends was sometimes sensitive to the fact that any potential repair solution, rather than an acceptable solution, was forthcoming. In the vast majority of cases, however, the

---

[4]For Cha'palaa there were only 40 total OIR cases with holds, so all cases were selected.

[5]We excluded elements like in-breaths and clicks from turn onset, because we could not always reliably observe their presence or absence in the recordings or judge whether or not interlocutors were able to perceive them.
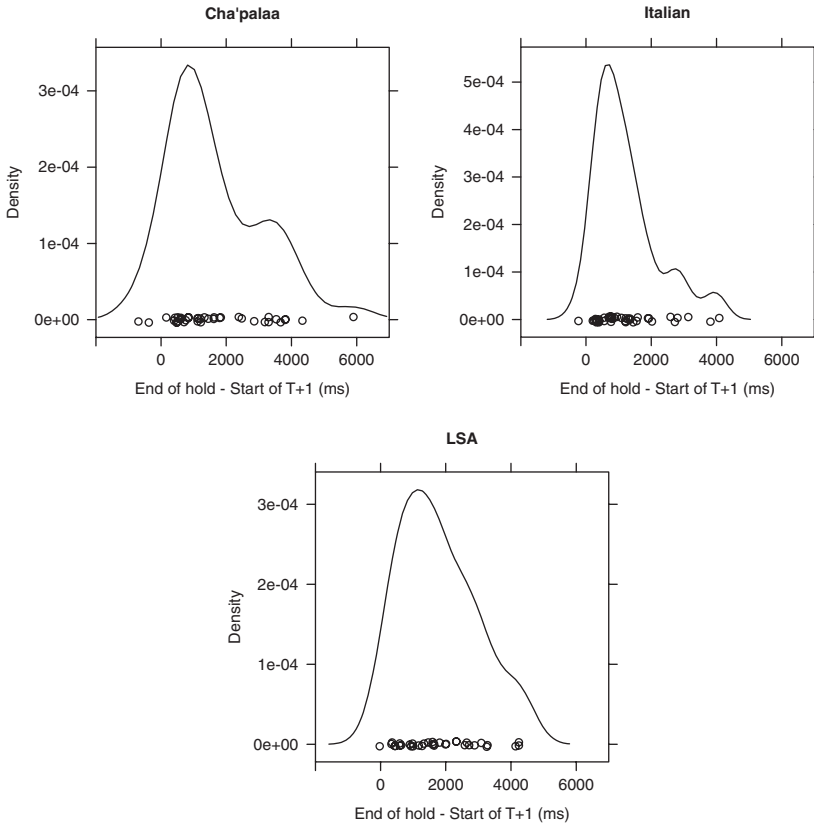
FIGURE 1   Density plots of the timing of hold ends relative to the start of T + 1 in 120 OIR sequences: 40 each from Cha'palaa, Italian, and LSA. [Density plots display the estimated probability density function (*y*-axis) of a continuous random variable (*x*-axis) and have a purpose similar to that of histograms. However, whereas histograms group observations into a discrete number of bins, density plots provide a continuous estimate of the distribution of a variable. The density plots shown in this article were computed using the density function in R with default parameter settings (R Development Core Team, 2008).]

hold was disengaged only once B had heard or had seen some or all of the repair solution and could thus judge whether it acceptably solved the problem that had prompted repair initiation. In general, then, when the parties who initiate repair accompany their spoken or signed T0 with some held bodily configuration, there is a strong tendency to disengage this configuration only when there is evidence that a repair solution is forthcoming.

Figure 1 shows that hold disengagement in repair sequences overwhelmingly occurs only after the resolution of the repair sequence is under way. When B stops

progressivity in T0, he or she makes relevant a response in the form of a repair solution on the part of A. We interpret B's maintenance of bodily configuration after T0 as a display to A that he or she is still committed to the sequence and expects the repair solution, whereas disengaging from this position displays that he or she no longer expects a solution because one has been provided. In the sequences that concluded at $T + 1$, this display is often the only immediate indication by B that A's repair solution was acceptable. However, sequences that include a $T + 2$ feature both a spoken or signed sequence-closing turn and a hold disengagement by the same participant.

Figure 2 shows density plots (with one removed outlier for Italian) of the timing of hold ends relative to the start of $T + 2$ (information uptake or sequence continuation by B) and shows the alignment of hold disengagement with $T + 2$ onset is more tight than that of hold disengagement to $T + 1$ onset (compare with Figure 1, which uses the same scale on the *x*-axis). Part of the reason for this is that B must often wait until A has produced part or all of $T + 1$ to ratify it as an acceptable solution, whereas at other times this can be projected earlier. In sequences with $T + 2$, on the other hand, the sequence-closing turn and the hold disengagement are produced by the same person (B) for the same reasons, and so these may be simultaneous or one may closely anticipate the other. The timings relative to $T + 1$ and $T + 2$ can be seen to complement each other, because they show how hold disengagement is timed after $T + 1$ onset and targeting $T + 2$ onset (in cases with $T + 2$). By definition, $T + 2$ must come after $T + 1$, and the slight delay in hold disengagement after $T + 1$ onset shows both how B waits until $T + 1$ is launched or complete to disengage the hold and how B often begins $T + 2$ around the same moment. However, it is worth noting some slight variation that speaks to issues of modality. Although for all three languages hold disengagement is tightly concentrated around the moment of $T + 2$ onset, for the spoken languages the visual bodily signal might occur slightly before or after the spoken elements of $T + 2$. On the other hand, for LSA, launching a new turn means encoding linguistic elements in the visual modality, which usually involves disengaging the current bodily position in order to resume signing.[6]

For LSA, then, hold disengagement was much more precisely aligned with $T + 2$ onset than for the spoken languages. Aside from this slight difference, which appears largely attributable to the articulatory facts of modality differences, the timing data show that the same process is at play in all three languages. Speakers

---

[6]Although there is a strong tendency for holds to be disengaged as signing resumes, this is not a given. Because the face, head, hands, and body provide a range of resources for simultaneous articulation, signers in some cases are able to maintain holding behavior while beginning to sign. However, we did not observe this very frequently.
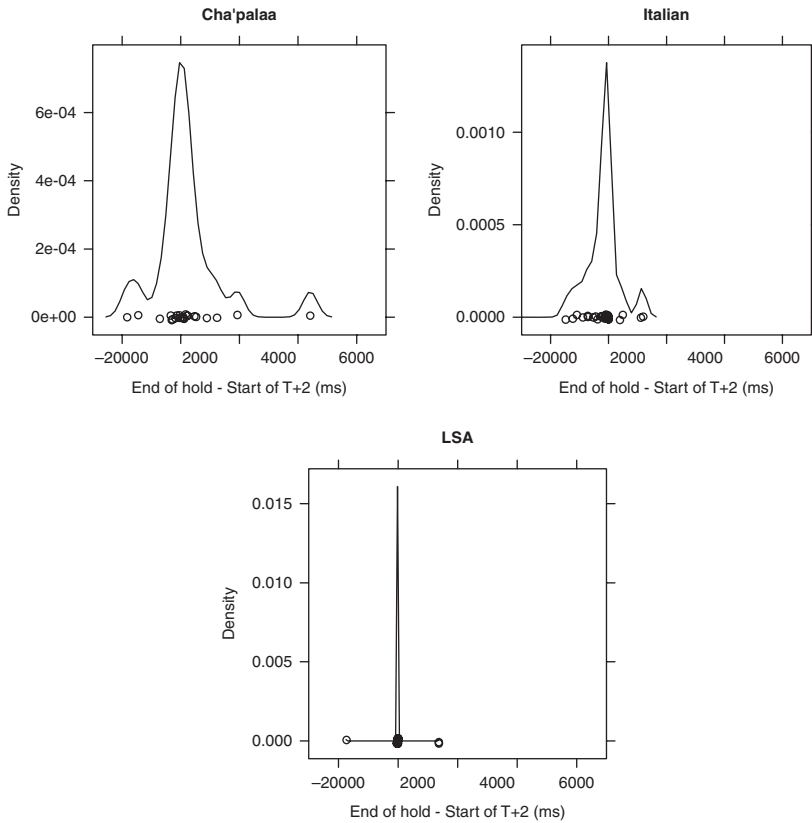
FIGURE 2   Density plots of the timing of hold ends relative to the start of T + 2 for a total of 95 of the 120 total cases; 28 in Cha'palaa, 32 in Italian (with 1 outlier not shown), and 34 in LSA.

and signers who initiate repair target their hold disengagement to the moment in the sequence when it is relevant to display that it has been successfully resolved.

Types of Visual Bodily Behavior

In addition to timing information, we coded each case for the type of visual bodily configuration that was held, including gaze, head direction (left/right), head position (up/down), leaning, eyebrow position, manual gesture, and signs (Figure 3). This wide range of visual behaviors, as well as any combinations of them, all fit with our definition of a hold, because they are all bodily configurations that can be assumed through some dynamic movement during T − 1, maintained in a static position during and after T0, and contrasted with a

Held behavior	Chapalaa (*n*=40)	Italian (*n*=40)	LSA (*n*=40)	All languages (*n*=120)
Gaze	95%	85%	97.5%	92.5%
Head re-direction (L/R)	82.5%	57.5%	27.5%	55.8% v
Head position (up/down)	5%	10%	62.5%	25.8%
Leaning upper body	20%	17.5%	32.5%	23.3%
Eyebrow position	7.5%	15%	92.5%	38.3%
Manual gesture	5%	5%	0%	3.3%
Manual sign	0%	0%	72.5%	24.2%

FIGURE 3   Frequencies of visual bodily resources held until OIR sequence closure.

dynamic release from this position after T + 1. These resources could be used individually, but in most cases more than one was present. They also showed variation across languages, some of which can be related to modality differences. For example, signers of LSA held manual signs 72.5% of the time, whereas this resource was obviously not available in spoken languages, which showed gestural manual holding but at a much lower rate (5%).

On the other hand, head redirection was much more common in the spoken languages (82.5% in Cha'palaa and 57.5% in Italian) relative to LSA (27.5%). This is because being engaged in sign language conversation requires that signers face each other much of the time, whereas speakers of spoken languages may change head direction more often without risking losing their perception of the speech signal. Egbert (1996) documents a pattern in German interaction in which specific repair initiators correlate with the re-establishment of mutual gaze during repair initiation; in spoken interaction, this type of movement into a static position around T0 complements the phenomenon of moving out of a static position that we focus on here. Although signers did not often move in and out of facing their interlocutor as observed in spoken language conversation, they could hold their head in upward or downward positions while still facing the addressee; this practice occurred often in LSA (62.5%) but relatively infrequently in Italian (10%) and Cha'palaa (5%). Held head position in LSA often occurs together with question-marking by raised or furrowed eyebrows, which were the most common nonmanual formats used to initiate repair (92.5%) (for more details on the LSA repair system see Manrique, in press).

In contrast with LSA, eyebrow position was held less frequently in Cha'palaa (7.5%) and Italian (15%). This is expected on the basis of the fact that OIR formats frequently include interrogatives, and eyebrow position is a grammaticalized marker of interrogativity in LSA (as in a number of other sign languages)

(Dachkovsky & Sandler, 2009; De Vos, van der Kooij, & Crasborn, 2009; Liddell & Johnson, 1989; Zeshan, 2004). Note, however, that the rates of the spoken languages may be conservatively low, because speakers' faces were not always clearly visible in the video, whereas in the LSA corpus they mostly were, in part due to the advantage of having two video recordings from cameras at different angles. However, it is fair to say that if in the spoken languages eyebrow position has a loose association with interrogativity (Ekman, 1979; Flecha-García, 2010), in LSA it is a grammatical and therefore normative practice.

Several of the visual resources we coded were used quite similarly in sign language and spoken languages. Of all the resources we studied in terms of holding practices, gaze was the most recurrent element of holds, occurring an average of 92.5% of the time. LSA predictably had the highest frequency, but there was minimal difference across the languages.[7] Another practice that was held at relatively similar frequencies across languages was leaning the upper body, which occurred 23.3% of the time in the three-language sample. In most of these cases the repair-initiating party leaned toward the addressee, displaying intensified attention and minimizing potential disruptions to the signal (see Rasmussen, 2014). In a few cases, particularly among signers, repair-initiators also leaned backward as a way to enable a wider visual channel, but whichever direction they leaned, both speakers and signers held this posture until sequence closure.

HOLDS IN PURSUIT SEQUENCES

The repair sequences analyzed in the previous section all feature some form of closure at or after T + 1. This means each sequence featured a successful repair solution by A and in many cases also a sequence-closing turn by B that ratifies the solution or resumes the progressivity of the halted sequence. However, in some cases the trouble was not resolved at the first T + 1 turn, and the sequence included further T0 turns, thus leading to an expanded sequential structure. As mentioned above, our timing measurements were based on sequence closure. Because, as we propose, visual bodily holds orient to the continued conditional relevance of the sequence, the longer the sequence remains unresolved the longer until the hold is disengaged. For example, in Extract 3 (line 6, Image 3A) Eva initiates repair once with the "open class" initiator *che* ('what?') (Drew, 1997), but when the repair solution introduces a new problem of pronoun ambiguity, she pursues an additional response, this time with a more specific "restricted-class" format, *chi* ('who?') (line 9, Image 3B).

---

[7]Although we did not measure the durations of time that gaze was diverted after hold disengagement, in general signers appear to return their gaze more quickly to continue the interaction.

(3)

Capodanno01_1813805

1 Ada ma lo sai che il Michele si legge i notturni di Chopin Eva

but 3S.A know-2S CMP the NAME 3S.D read-3S the nocturnes of Chopin NAME

do you know that Michele reads Chopin's nocturnes Eva?

2 Remo ho visto

have–1S see-PSTP

I saw that   T–1

3 (0.5)

4 Eva che: ((turns to Ada)) **wha:t?**	T0

5 (0.7)

6 Ada si legge dei notturni di Chopin

3S.D read-3S some nocturnes of Chopin

he reads Chop in nocturnes   T+1

7 (0.5)

8 Ada pian pian[ino con la] destra e la sinistra=

slowly slowly-DIM with the right and the left

little by little with the right and the left hand

9 Eva [chi] ((keeps looking at Ada)) **who?**	T0

10 Ada =il Michele   T+1

the NAME

Michele

11 Eva ma dai ((looks away from Ada and down on the worktop)) T+2

but PCL

no way!

IMAGE 3A Line 4: 'what?'. Eva (on the left) turns her head and initiates repair. Hold begins.

IMAGE 3B   Line 9: 'who?'. Hold is maintained.

IMAGE 3C   Line 11: 'no way'. Eva looks away. Hold is released.

Pursuits in repair sequences commonly show the same ordering of repair initiator formats, with an open-class initiator as the first T0 and a restricted-class initiator as the second (and often final) T0 (Clark & Schaefer, 1987, p. 23; Enfield et al., unpublished data; Schegloff et al., 1977, p. 369; Svennevig, 2008). The structure of Excerpt (4) from LSA is closely parallel to that of (3) from Italian (an identically structured "huh?/who?" case was found in the Cha'palaa corpus as well). In Extract 4, participant B initiates an open repair first (line 2, Image 4A), followed by a restricted initiator ('who?') in line 4 (Image 4B). The composite utterance in the first repair initiation consists of a manual sign 'WHAT?' together with nonmanual interrogative markers: eyebrows held together intensified with a wrinkled nose. At the end of line 2,

participant B maintains the eyebrows in a stationary position while beginning to pursue the solution, first puckering her lips, then producing the question word '*WHO?*' (while also mouthing the Spanish question word 'who': '*quién*'), and then holding this sign together with the nonmanual markers (eyebrows together, head tilt; line 4) until the repair solution in line 5.

(4) Ayrolo_843376

| 1 | A | WELL MOTORCYCLE-ONE TWO-CARS | ←T–1 |
| | | **Well,   we have one motorcycle and two cars** | |

2		Q-ET-------------**H** (hold)	
		Q-NW---------**H PLT** (puckered-lips-forward)	
		BWHAT--------------**H** (1.7)	←T0
		que (mouthed Spanish wh-q word)	
		What?	

| 3 | C | MOTO THERE-IS-NOT CAR | ←T+1 |
| | | **There is no motorbike, only car** | |

4	B	____Q-ET	
		WHO------- [-----**H**(0.9)	←T0
		Who?	

| 5 | C | [SN-H [SN-H (1.3) ((looking at A)) | ←T+1 |
| | | **John, John,** | |

| 6 | B | [*H-Up* | ←T+2 |
| | | **Ah !** | |

IMAGE 4A   '*WHAT?*', participant B on the right initiates an open repair type by the manual sign '*WHAT*', bringing her eyebrows together and wrinkling her nose. These linguistic lexical and prosodic components are maintained until participant C, middle, provides a repair solution Extract 4, line 2.

IMAGE 4B   '*WHO?*', participant B, right, upgrades her repair initiation format to restricted type produced by the manual sign '*WHO*' combined with eyebrows together, head tilt and mouthing of the Spanish wh-q word in '*quién*', Extract 4, line 4. All these linguistic components are produced simultaneously.

IMAGE 4C   '*Ah*!', participant B provides uptake by nodding and by releasing her hold posture by lowering her right arm, Extract 4, line 6.

Complex repair sequences featuring pursuits show an analogous use of holds across languages: when the sequence expands, the hold is prolonged. In some instances one element of a hold may be released earlier in the sequence, whereas another is maintained to the end, as in some cases when A lowered her or his hands while maintaining gaze and head direction, but the overall tendency was to release all elements together. By comparing simple and expanded sequences we can see how the visual bodily behavior orients directly to the unfolding contingencies of interaction, with the duration of the holds lengthening together with the sequence duration.

SEQUENCES WITHOUT HOLDS

The data presented above represents a complex of interactive practices in which the speakers of first pair parts recurrently hold and disengage bodily configurations to display their orientations to unresolved or resolved status of the repair sequence they are involved in. But this practice is not obligatory for initiating and closing

repair sequences (indeed, repair sequences are frequent in phone calls too). One reason why it might not occur is that mutual gaze may be obstructed or interactants may be engaged in some other physical activity as they converse, in which case they have a ready account for why they may not be expected to display their orientation to sequence closure through bodily behavior. For example, in the following excerpt from Cha'palaa, a woman is weaving a basket as she talks. She initiates and closes a repair sequence without any relevant change in her body position while her gaze remains fixed on her hands and the basket fibers.

(5) CHSF2011_06_25S2_859221

1 A   wee-mujtu-sha, santa ma- santa mariya-sha
        other-place-in   santa ma- santa maria-in
        in other places, in Santa Maria   ←T–1

2   ma dolar medio-chi kan-tsa-a   de-tin-ti-ee   yumaa
        one dollar half-with get-PROG-FOC PL-say-say-N.EGO now
        they say they are buying them at a dollar and a half

3 B   pute-e
        basket-Q
        baskets?   ←T0

4 A   jee
        yes   ←T+1

5 B   aa
        oh   ←T+2

((B is making a basket and does not look up))

IMAGE 5   'baskets?', speaker B on the right initiates repair on A, in the center; B is occupied with weaving a basket and does not display or hold any relevant bodily behavior as A confirms: "yes".

In cases like (5), participants orient to the concurrent ongoing activity by adapting their interactive practices such that less use is made of visual bodily resources and presumably less intense visual "mutual monitoring" (Goodwin, 1980). In other cases holds may be begun but not sustained, due to some contingencies that arise during the sequence. In Excerpt (6) participants A, B, and C are talking and at the same time are busy with practical activities involved in styling A's hair. A asks B to go get some more paper towels "so he'll always be ready" and as B walks away he initiates repair (line 7, Image 6A) to disambiguate what he should be ready for. While walking to the kitchen counter, he torques his upper body to gaze at A and maintains this configuration past the end of T0. However, as his body nears the counter, he breaks the hold and focuses his attention on getting the paper towel, releasing the hold before A begins producing T + 1.

```
(6) Tinta 329287

1 C        ti:: ti  serve  mi  sa      questo ((holds out kitchen paper))
           2S.D 2S.D serve-3S  1S.D know-3S this
           you:: you  need  this  I  feel

2          (1.5)

3 B        sì adesso mi  serve
           yes now     1S.D serve-3S
           yes now  I  need  it

4 A        ecco prendine        dell'altro che è    là
           PCL  take-NPST-2S=PRT PRT=other  RELbe.3S there
           right get  some  more  from  there

5          così almeno tu   sei  sempre pronto                  ←T–1
           so   at.least 2S.N be.2S always  ready
           so  at  least  you'll  be  always  ready

6          (0.8)

7 B        a soffiare il  naso                                  ←T0
           to blow-INF  the nose
           to  blow  the  nose?

8          (0.2)

9 A        sì                                                   ←T+1
           yes
```

For the Italian and Cha'palaa corpora, which include many practical household activities, no relevant visual holding behavior was observed in approximately 40% of cases. In the LSA corpus, on the other hand, we did not find sequences comparable with the Cha'palaa and Italian excerpts above, because almost all repair sequences had some kind of holding behavior maintained at least until the beginning of T + 1. This may partially be explained by the fact that in sign language conversation, it is not as easy to maintain a concurrent physical activity that diverts visual attention; additionally, in the LSA corpus conversation was generally the main activity, whereas in the other corpora interactants engaged more frequently in parallel activities. The circumstances under which holds did

IMAGE 6A   Line 7: "To blow the nose?". B initiates repair and briefly holds his head direction.

IMAGE 6B   Line 8: B disengages head position before A's confirmation ("yes") in line 9.

not occur in OIR sequences were varied, but they were primarily related to practical activities that required attention or to instances in which the bodily configurations of the participants complicated mutual visual perception. Such instances were common in the spoken language corpora, and we take the lack of holds in these cases to be evidence of the flexibility of the repair system. The fact that a small percentage of LSA cases also featured no discernable hold indicates that some instances of lack of holds in OIR sequences may also be interactionally meaningful independently of complicating circumstances in ways that further study may uncover, both in signed and spoken languages.

DISCUSSION

In this article we asked about the role of visual bodily behavior in repair sequences, focusing on moments when the hands, head, eyes, and upper body are held stationary for periods of time. What we find is that when such holds occur in repair sequences, their timing overwhelmingly conforms to the following generalization: when some aspect of B's visual bodily behavior is held through the end of T0, it is held by B at least until the onset of the repair solution (T + 1) and is then disengaged during or shortly after this turn, around the onset of the sequence-closing next turn (T + 2), if there was one. We interpret this finding to mean that holds are a cross-linguistic resource available to both speakers and signers for closing repair sequences, and that holds are regularly produced—and probably normatively expected—under most circumstances of face-to-face interaction in which participants are not engaged in competing activities that require dividing their attention or moving out of a configuration in which gaze and bodies are oriented towards each other ("F-formation," Kendon, 1990). These conditions vary between signed and spoken languages, because speakers may orient their body away from the interaction with less risk of stopping the conversation than signers can. Some of the specific resources that speakers and signers tended to hold vary as well, but whatever the hold's composition, it conformed to the same principles of timing regardless of modality.

This result speaks to broader questions about conversational turn-taking in multimodal, face-to-face interaction. It shows that elements of the broader composite utterance that are not part of the turn-taking system proper are nonetheless closely timed to turn-based sequential structures. Visual bodily behaviors are organized on a parallel system that allows overlap with the turns of other participants. In cases like those seen above, holds were maintained in overlap across sequences of three or four turns. This has been shown for gaze and other nonverbal behaviors in sequences other than repair (e.g., Rossano, 2012). For sign languages, there is evidence to suggest that although signers tend to minimize overlap in strokes, this does not necessarily apply to other phases, including the phase of poststroke holds, which can overlap with subsequent turns (De Vos, Torreira, & Levinson, 2015; Kita, 1990; McNeill, 1992). Additionally, our results suggest that this parallel interaction track is not simply determined by the turn-based sequence but that it can affect its course: in absence of a T + 2 turn, hold disengagement is the most immediate signal that an OIR sequence has been resolved.

For methodological reasons it was important at the beginning of our study to treat the spoken elements of repair sequences and the timing of visual bodily holds as independent phenomena. However, we are now in a position to say something about their close functional relation. Indeed, the best account is not simply that

hold disengagement *accompanies* sequence closure but that it concurs to *accomplish* it, either in combination with a spoken turn or, in cases where there is no spoken or signed T + 2, acting as a "visual T + 2." Revisiting the data presented above, this helps to see how the sequences with no T + 2 were resolved.

In addition to their function in repair sequences, visual bodily practices like those we examined appear to be broadly relevant for a wide range of interactive practices when different types of conditional relevance is at play. In our corpora we informally observe similar holding practices timed to sequence closure in question–answer sequences and request sequences, as two other contexts in which speakers may display ongoing commitment to resolving ongoing sequences. Like in OIR sequences, in these contexts this resource has a special affordance for pursuing a response because it can be displayed such that the expected second pair part can begin without problems of overlap. We expect that further research on a wider range of sequence types will reveal a family of holding practices with analogous functions.

Finally, the most compelling question raised by these data is *why* similar timings of holds relative to similar sequence types can be observed in conversational interaction across radically different languages. These facts are in alignment with an emerging body of evidence indicating that whereas grammatical systems show almost endless diversity (Evans & Levinson, 2009), basic interactive practices show much less variance cross-linguistically (Enfield & Levinson, 2006; Enfield et al., 2013; Levinson, 2000; Stivers et al., 2009). The reasons for this are complex and difficult to pinpoint, but we find that the best account concerns similar material, temporal, and social conditions in the face-to-face speech situation. As a basic prerequisite for interaction, speakers must be able to perceive each other's turns before reacting to them, and this generates some basic constraints of temporality and bodily orientation that apply cross-linguistically. Because of their connection to the attentional dynamics of interacting in physical bodies, elements like gaze and posture are provided with natural meaning that can additionally function as semiotic displays of engagement (cf. Enfield, 2009; Enfield et al., 2007; Grice, 1957).

Dingemanse, Enfield, and Torreira (2013) explain another formal aspect of repair sequences—the form of the open-class repair initiating interjection like "huh" in English—across diverse languages through processes of convergent evolution, in which interactants under similar conditions converge on similar solutions. Relative to the spoken elements of repair sequences, visual bodily practices are under their own specific pressures, and despite their stark contrasts in terms of grammatical systems and modality, speakers of LSA, Italian, and Cha'palaa use similar visual bodily practices in similar sequential contexts. We predict that adding more languages to this small yet diverse sample will yield results that pattern the same way, because in any language speakers must confront the same basic social and physical pressures of interaction. The fact that

other researchers have described related practices in several languages makes us confident that the interactional patterns of visual bodily behavior we observed in our comparison are likely to be broadly relevant for the study of multimodal interaction cross-linguistically, rooted in the similar pressures and affordances of face-to-face interaction across social contexts.

ACKNOWLEDGMENTS

REFERENCES

Baker, C. (1977). Regulators and turn-taking in American Sign Language discourse. In L. Friedman (Ed.), *On the other hand: New perspectives on American Sign Language* (pp. 215–236). New York, NY: Academic Press.

Bolden, G. B., Mandelbaum, J., & Wilkinson, S. (2012). Pursuing a response by repairing an indexical reference. *Research on Language & Social Interaction*, *45*, 137–155.

Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies*, *7*, 507–525.

Clark, H. H., & Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, *2*, 19–41.

Dachkovsky, S., & Sandler, W. (2009). Visual intonation in the prosody of a sign language. *Language and Speech*, *52*, 287–314.

De Vos, C., Torreira, F., & Levinson, S. C. (2015). Turn-timing in signed conversations: Coordinating stroke-to-stroke turn boundaries. *Frontiers in Psychology*, *6*, 268. doi:10.3389/fpsyg.2015.00268.

De Vos, C., van der Kooij, E., & Crasborn, O. (2009). Mixed signals: Combining linguistic and affective functions of eyebrows in questions in sign language of the Netherlands. *Language and Speech*, *52*, 315–339.

Dingemanse, M., & Floyd, S. (2014). Conversation across cultures. In N. J. Enfield, P. Kockelman, & J. Sidnell (Eds.), *The Cambridge handbook of linguistic anthropology* (pp. 447–480). Cambridge, UK: Cambridge University Press.

Dingemanse, M., Kendrick, K. H., & Enfield, N. J. (in press). A coding scheme for other-initiated repair across languages. *Open Linguistics*.

Dingemanse, M., Torreira, F., & Enfield, N. J. (2013). Is "huh?" a universal word? Conversational infrastructure and the convergent evolution of linguistic items. *PLoS ONE*, *8*, e78273.

Drew, P. (1997). "Open" class repair initiators in response to sequential sources of troubles in conversation. *Journal of Pragmatics*, *28*, 69–101.

Egbert, M. M. (1996). Context-sensitivity in conversation: Eye gaze and the german repair initiator bitte? *Language in Society*, *25*, 587–612.

Ekman, P. (1979). About brow: Emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology: Claims and limits of a new discipline* (pp. 169–202). Cambridge, UK: Cambridge University Press.

Enfield, N. J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge, UK: Cambridge University Press.

Enfield, N. J., Dingemanse, M., Baranova, J., Blythe, J., Brown, P., Dirksmeyer, T., & … Torreira, F. (2013). Huh? What? A first survey in 21 languages. In M. Hayashi, G. Raymond, & J. Sidnell (Eds.), *Conversational repair and human understanding* (pp. 343–380). Cambridge, UK: Cambridge University Press.

Enfield, N. J., Kita, S., & de Ruiter, J. P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*, *39*, 1722–1741.

Enfield, N. J., & Levinson, S. C. (Eds.). (2006). *Roots of human sociality: Culture, cognition, and human interaction*. Oxford, UK: Berg.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*, 429–448; discussion 448–494.

Flecha-García, M. L. (2010). Eyebrow raises in dialogue and their relation to discourse structure, utterance function and pitch accents in English. *Speech Communication*, *52*, 542–554.

Goffman, E. (1981). *Forms of talk*. Philadelphia, PA: University of Pennsylvania Press.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York, NY: Academic Press.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, *32*, 1489–1522.

Goodwin, M. H. (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry*, *50*, 303–317.

Grice, H. P. (1957). Meaning. *Philosophical Review*, *66*, 377–388.

Groeber, S., & Pochon-Berger, E. (2014). Turns and turn-taking in sign language interaction: A study of turn-final holds. *Journal of Pragmatics*, *65*, 121–136.

Hayashi, M., Raymond, G., & Sidnell, J. (Eds.). (2013). *Conversational repair and human understanding*. Cambridge, UK: Cambridge University Press.

Johnston, T. (2010). From archive to corpus: Transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics*, *15*, 106–131.

Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge, UK: Cambridge University Press.

Kita, S. (1990). The temporal relationship between gesture and speech: A study of Japanese-English bilinguals. Unpublished master's thesis. University of Chicago, Chicago, Illinois.

Kita, S., Gijn, I. van, & Hulst, H. van der (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction* (pp. 23–35). New York, NY: Springer.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Levinson, S. C. (2013). Action formation and ascription. In T. Stivers & J. Sidnell (Eds.), *Handbook of conversation analysis* (pp. 103–130). Malden, MA: Wiley-Blackwell.

Liddell, S. K., & Johnson, R. E. (1989). American Sign Language: The phonological base. *Sign Language Studies*, *1064*, 195–277.

Li, X. (2014). Leaning and recipient intervening questions in Mandarin conversation. *Journal of Pragmatics*, *67*, 34–60.

Manrique, E. (in press). Other-initiated repair in Argentine Sign Language. *Open Linguistics*.

Manrique, E. (2011, May). Other-repair initiators in Argentine Sign Language: Handling seeing and understanding difficulties in face-to-face interaction. Unpublished master's thesis. Radboud University Nijmegen, Nijmegen, The Netherlands.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

Mondada, L. (2006). Participants' online analysis and multimodal practices: projecting the end of the turn and the closing of the sequence. *Discourse Studies*, *8*, 117–129.

Mondada, L. (2007). Multimodal resources for turn-taking pointing and the emergence of possible next speakers. *Discourse Studies*, *9*, 194–225.

Park-Doob, M. A. (2010). Gesturing through time: Holds and intermodal timing in the stream of speech. Unpublished doctoral dissertation. University of California, Berkeley, Berkeley, California.

Pomerantz, A. (1984). Pursuing a response. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 152–163). Cambridge, UK: Cambridge University Press.

Rasmussen, G. (2014). Inclined to better understanding—The coordination of talk and "leaning forward" in doing repair. *Journal of Pragmatics*, *65*, 30–45.

Rossano, F. (2012). Gaze behavior in face-to-face interaction. Unpublished doctral dissertation. Radboud University Nijmegen, Nijmegen, The Netherlands.

Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning and culture. In J. Sidnell (Ed.), *Conversation analysis: Comparative perspectives* (pp. 187–249). Cambridge, UK: Cambridge University Press.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696–735.

Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, *70*, 1075–1095.

Schegloff, E. A. (1984). On some questions and ambiguities in conversation. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 28–52). Cambridge, UK: Cambridge University Press.

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, *97*, 1295–1345.

Schegloff, E. A. (1998). Body torque. *Social Research*, *65*, 535–596.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge, UK: Cambridge University Press.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*, 361–382.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*, 289–327.

Sikveland, R. O., & Ogden, R. A. (2012). Holding gestures across turns: Moments to generate shared understanding. *Gesture*, *12*, 167–200.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., & ... Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*, 10587–10592.

Stivers, T., & Rossano, F. (2010). Mobilizing response. *Research on Language & Social Interaction*, *43*, 3–31.

Svennevig, J. (2008). Trying the easiest solution first in other-initiation of repair. *Journal of Pragmatics*, *40*, 333–348.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). *ELAN: A professional framework for multimodality research*. Retrieved from http://tla.mpi.nl/tools/tla-tools/elan/

Zeshan, U. (2004). Interrogative constructions in signed languages: Crosslinguistic perspectives. *Language*, *80*, 7–39.

APPENDIX

Organization of the Lines of the Transcripts and Conventions for Glosses in LSA

	Q-ET------------h (hold)	⎫
	Q-NW--------h_PLT (puckered-lips-forward)	⎬ nonmanual signs
B	WHAT---------------h (1.7)	⎭ ⟶ manual sign
	Quién (Spanish wh-q word for *who*)	⟶ mouthed words
	What?	⟶ free English translation

CAPITAL LETTERS = sign glosses

ER = eyebrows raised

ET = eyebrows together

F-I-N-G-E-R-S-P-E-L-L-I-N-G = finger spelling is indicated by hyphen between letters

G: gesture = followed by a description of the meaning of the gesture, e.g., G:I-don't-mind

H = hold

h-down = head down

h-up = head up

HYPHENATED-WORDS = represent a single sign and more than one English word

NEGATIVE-VERB = glossed with the negation in a post verb position, e.g., KNOW-NOT

NW = nose wrinkled

PLT = puckered lips forward

PRO1 = PRO: pronoun, 1: first person, 2: second person, 3: third person

SN = sign name

Key to Glosses in Spoken Languages

Cha'palaa: Q = question, FOC = focus, PL = plural, PROG = progressive aspect, N.EGO = non-egophoric (speaker did not instigate event).
Italian: 1 = first person, 2 = second person, 3 = third person, A = accusative, CMP = complementizer, D = dative, DIM = diminutive, INF = infinitive, N = - nominative, NAME = proper name, NPST = Non-Past, P = plural, PCL = particle, PRT = partitive, PSTP = Past Participle, REL = relativizer, S = singular.
The unmarked verb inflection is present indicative (simple present).

Conversation Analytic Conventions

- Visual bodily behavior is described in double parenthesis, e.g.:
  ((turns away from A))
- Conversational overlap is represented with square brackets, e.g.:
  A: pian pian[ino con la] destra e la sinistra =
  B: [chi   ]
- Vowel extension is indicated with a colon, e.g.: che:
- Pause length is indicated in single parenthesis in units of seconds, e.g.: (1.5)