The face of communication breakdown: Multimodal repair in L2 oral proficiency interviews

J. Dylan Burton
Michigan State University

All communication entails moments of breakdown and repair. Research has shown that test-takers in oral proficiency interviews use both verbal and nonverbal behaviors to navigate these breakdowns, which may occur more often in online speaking tests than face-to-face settings. While individuals are often idiosyncratic in their baseline nonverbal behavior, it is unknown to what extent speakers resemble or differ from each other during moments of repair. To date, a research agenda including a broader range of embodied behavior in the context of language tests is lacking. In this study, four Zoom-based video recordings of second language tests were used to investigate seven categories of behavior and their alignment with speech during moments of other-initiated repair., This contribution uses multimodal conversation analysis to analyze differences in behavior amongst the participants and how those patterns align with the turn-by-turn nature of repair. The analysis shows that participants varied substantially in both the frequency and duration of each type of behavior during repair, and furthermore differed in how they combined these during their interaction with the examiner. The study highlights the need for more research in how individuals demonstrate their interactional competence both verbally and nonverbally for greater construct representation in second language speaking tests.

Keywords: NVB, repair, miscommunication, speaking assessment, language testing

Language is used so that people can share their thoughts and ideas. Broadly speaking, when people speak to each other, they are engaged in idea sharing, but this flow of this exchange of ideas can break down when one or more speakers realize that they did not understand a previous utterance due to linguistic or acoustic problems. In these instances, listeners may attempt to initiate a repair to the conversation, or they can skip it and carry on. In second language (L2) oral proficiency interviews (OPIs), these comprehension

Email address for correspondence: burtonjd@msu.edu

problems form an important part of the fabric of test discourse due to inherent linguistic and cognitive demands on the test-taker. Although these interactions are primarily verbal, as is the case of all discourse they may also include a wide range of semiotic resources including nonverbal behavior (NVB).

In order to communicate in paired or interview-based language tests and repair these breakdowns in communication, L2 speakers draw on interactional competence (IC), or "the ability to co-construct interaction in a purposeful and meaningful way, taking into account sociocultural and pragmatic dimensions of the speech situation and event" (Galaczi & Taylor, 2018, p.226). IC has been conceptualized as distinct from language proficiency in that IC is socially situated and context dependent, but language proficiency is primarily a psycholinguistic attribute and thus context independent (Roever & Dai, 2021). IC, like language proficiency, is multi-componential and may be "supported by the linguistic and other resources that speakers and listeners leverage at a microlevel of the interaction, namely, aspects of *topic management, turn management, interactive listening, breakdown repair and non-verbal or visual behaviours* (Galaczi & Taylor, 2018, p. 226, emphasis in original). The assessment of IC has received increasing attention in recent years through treatment in special issues of language assessment journals and edited volumes dedicated to the topic (Salaberry & Burch, 2021; Salaberry & Kunitz, 2019; Plough et al., 2018; Youn & Burch, 2020). Nonetheless, Galaczi and Taylor (2018) noted in their conceptualization of IC that some facets of interaction, such as NVB, have received unequal empirical study. Plough (2021) lamented this absence of research, stating that:

> The speaking construct in face-to-face interaction is underrepresented if NVB is excluded… Perhaps a unique characteristic of NVB that makes it particularly elusive is its fluidity and contextual variation – individual variation in the interpretations of the same NVB as well as individual variation in the production of NVB (p. 62).

Ensuring NVB has a place in the speaking construct can only be achieved through studies documenting its role in various functions and contexts, and thus descriptive studies relating NVB to other core IC competences (e.g., turn management or repair) are needed.

Studies have additionally begun to show that clarification request sequences, which are one type of repair after a potential breakdown in communication, tend to appear more frequently in online, video-conferencing-mediated L2 OPIs than in face-to-face tests (Cooke, 2015; Nakatsuhara, Inoue, Berry, & Galaczi., 2021), thus prompting the need for more research in this particular medium. In this paper, I will attempt to bridge these gaps by systematically addressing NVB within a multimodal conversation analytic (CA) framework. I will describe the use of a dense, micro-level multimodal transcription

conducted in ELAN software to document the unfolding comprehension problems of four test-takers taking an online OPI. My key aim is to detail both the communalities of their behavioral features as well as how individuals may differ within this narrow context, thus providing a descriptive account of the relationship between NVB and test-related repair, and documentation of the individual variation in this behavior.

Background

Oral proficiency interviews

In L2 educational contexts, OPIs are ubiquitous. They typically involve one test-taker and one examiner and may include role plays or monologic discourse, but their prototypical format is a question-answer response pattern: The examiner asks a question and the test-taker gives an answer. Generally, the examiner provides no scaffolding or extension and may be restricted in their ability to support comprehension breakdowns. Thus, OPIs are goal-driven, institutional phenomena that involve interactants who may experience very different perceptions of the unfolding event (He & Young, 1998) and thus undergo breakdowns in mutual understanding or *intersubjectivity*—"the sharing of experiential content (e.g., feelings, perceptions, thoughts and linguistic meanings) among a plurality of subjects" (Zlatev et al., 2008, p.1). This contrasts with general conversation, in which the goals and purposes of interactions may be less restricted and more unpredictable (Lazaraton, 2002).

The interactional patterns of OPIs may differ from conversational and classroom discourse (He & Young, 1998; Lazaraton, 2002; Ross, 2017; Seedhouse, 2013; Seedhouse & Egbert, 2006). Seedhouse (2013) found that the organization of interaction in OPIs differs from other institutional discourse in that 1) intersubjectivity is not required for both parties; 2) there is little to no feedback provided by the examiner; 3) topic development is often stymied; and 4) the interaction is asymmetrical in both how the examiner controls turns and has advance knowledge of upcoming questions. OPIs consist largely of pre-allocated turns as well as pre-allocated (or even pre-scripted) clarification sequences for the examiner—that is, comprehension problems in OPIs and examinee-initiated repair sequences are anticipated.

Other-initiated repair

Repair sequences may occur after communication between two individuals breaks down due to comprehension problems (Sacks et al., 1974). In OPIs, test-takers may fail to understand the examiner and then proceed to pursue repair by requesting repetition of

the test question or clarifications; in other words, the repair is *other-initiated* (Schegloff, 2000) by the listener-test-taker, and the examiner may proceed to repair the test-taker's breakdown in comprehension. This is differentiated from self-repair—"correction by the speaker of that which is being corrected" (Schegloff et al., 1977, p. 361)—whereby a speaker identifies a problem during speech production and corrects their own utterances through word searches, paraphrasing, recycling, insertions, and other actions. Other-initiated repair typically immediately follows the examiner's questions (Kim & Suh, 1998; Seedhouse & Egbert, 2006). Repair sequences generally contain four key elements: the *trouble-source*, the *initiation*, the *repair*, and the *resolution* (Benjamin & Mazeland, 2012). The trouble-source is the word or utterance that causes a comprehension difficulty, uttered by the first speaker (in this case, the examiner). The repair initiation is produced by the listener (the test-taker here) to signal a breakdown or failure in understanding to the speaker. The initiation may be open-class (e.g., 'Sorry?'; Drew 1997) or alternatively targeting the trouble-source with a restricted repair initiation (e.g., 'Did you say X?'; Dingemanse et al., 2016). After initiation, the original speaker then issues the repair, which may be a direct repetition of the first utterance or a reformulation or clarification of a word or entire phrase. If the repair is successful, this results in a resolution and communication continues. If the repair is unsuccessful, the sequence may repeat again until resolution is reached.

Recent research has found that clarification requests by the test-taker tend to occur more frequently in online OPIs than face-to-face (Cooke, 2015; Nakatsuhara, Inoue, Berry, & Galaczi., 2021). Nakatsuhara, Inoue, Berry, & Galaczi (2021) compared 30 test-takers' IELTS speaking test performances recorded in both online and face-to-face modes. After coding these files for a wide range of discourse and interactional functions, they found that the only difference in the use of functions between the two modes was the use of confirmation requests by the test-taker, occurring 36.6% more often in the online mode of the interview portion of the test. If these types of other-initiated repair indeed occur more often in the online mode, research is needed to describe and document the unfolding of these sequences in this increasingly common context.

There is a substantial body of research on other-initiated repair in L2 speech (Gardner & Wagner, 2004; Hellermann, 2011), some of which has also considered the role of NVBs (Floyd et al., 2016; McDonough et al., 2019; Mortensen, 2012; Oloff, 2018; Seo & Koshik, 2010). Repair has also played an important role in research on OPIs (Kim & Suh, 1998; Kim & Park, 2015; Roever & Kasper, 2018). Roever & Kasper (2018) noted "the need for the assessment of talk to consider other vocal resources than language … [test-takers] regularly mobilize a range of semiotic resources (e.g., gestures, gaze, body position and movement, objects and space) to accomplish the activity at hand" (p. 349). In order to

more fully understand the IC individuals use in OPIs or other language testing settings, it is critical to take into account test-takers' NVB as well (Plough et al., 2018; Plough, 2021).

NVB and repair in language tests

NVB, in particular facial expressions, serves a number of expressive, communicative, and regulatory roles (Ekman & Friesen, 1969), and may even "betray the character of a person, as well as their current affective state and intentions" (Kappas et al., 2013, p. 132). A number of studies to date have shown that NVBs such as gestures, gaze behavior, posture, facial expressions, and other paralinguistic cues are taken into account by rater-listeners in their perceptions of IC (Ducasse & Brown, 2009; May, 2011), and may even impact test-takers' final test scores (Jenkins & Parra, 2003; Gan & Davison, 2011; Nakatsuhara, Inoue, & Taylor., 2021). Overall, these studies have found that behaviors that contribute to an impression of expressiveness, confidence, and attention likewise positively impact perceptions of IC and/or language proficiency; on the other hand, static, less attentive, "stoic" behaviors may have a corresponding negative effect.

A second line of research in language testing, generally using CA, has considered the ways test-takers navigate the testing context using both verbal and nonverbal resources. The bulk of these have analyzed paired and group speaking tests (Burch & Kley, 2020; Greer & Potter, 2008; Hırçın Çoban & Sert, 2020; Nakatsuhara, 2011; van Compernolle, 2013). These studies have largely found that test-takers use facial expressions (e.g., eyebrow raising, smiling), averted/mutual gaze, posture, and gesture to manage turn-taking sequences (Greer & Potter, 2008; Nakatsuhara, 2011), to flag interactional trouble (Hırçın Çoban & Sert, 2020), and to convey intersubjectivity or shared understanding (Burch & Kley, 2020). Jenkins and Parra (2003) and Tominaga (2013), on the other hand, looked into multimodal cues in the unfolding of interaction in OPIs. Tominaga (2013) discussed the ways test-takers negotiate turn-taking with examiners, finding multiple functions of gaze (e.g., to hold the floor, complete turns, and indicate a lexical search) and eyebrow furrowing during lexical trouble and self-repair. Jenkins and Parra (2003) drew from an extensive literature in nonverbal communication to document the NVBs of more and less successful test-takers. They found that backchannels/receipt tokens (e.g., head nodding), displays of affect (e.g., laughing, smiling), and regulators (e.g., forward leaning, eye contact) were characteristic of more successful OPI interactions. Features that negatively impacted OPI performance included extended silence and a relative absence of head nodding, stiff posture, gaze aversion, and flat prosody.

These studies generally analyzed NVB in a discourse or CA tradition. In these transcriptions, the principal focus was on language, pauses, and turns, but nonverbal cues were often transcribed only as they became salient, or only focusing on one or two

features. What is missing is a systematic transcription of a broader range of cues occurring within particular sequences. Dense, micro-level transcriptions can offer substantially more insight into the interactions among various behaviors to describe unfolding phenomena in a particular context. This is particularly appropriate in the case of repair sequences in OPIs, as these often stem directly from a test question, can correspond to specific lexical or grammatical trouble sources, and normally do not extend across numerous turns. These descriptions of contextualized language can then be useful for evidence-centered modifications to test constructs and rating scales, as well as in the selection of speech samples for benchmarking. With this in mind, I propose the following research questions:

- Which nonverbal resources do test-takers employ when initiating and resolving repair sequences?
- How do individuals vary in their embodied behaviors while negotiating repair sequences as a result of communication breakdowns on a foreign language oral proficiency test?

Method

Context of the study

The data for this study were drawn from a dataset of online speaking tests conducted in summer 2020. The tests took place in Zoom (www.zoom.us). The test was an OPI with one examiner and one test-taker. It consisted of ten questions in question-answer format. The ten questions were ordered by increasing complexity, and all participants were given the same test. As the sole examiner, I piloted a procedure with three participants prior to data collection to standardize my verbal and nonverbal reactions to the test-takers, as uncontrolled reactions or behavior can introduce variable influence on the participants' comprehension (Faraco & Kida, 2008). This standardized behavior included a) ensuring that the position of my head and body were centered and clearly lit, b) maintaining gaze with the test-taker during their responses to appear attentive and interested, c) limiting backchannels to the nonverbal mode (e.g., head nodding) and avoiding verbal backchannels to reduce interruptions in Zoom, and d) keeping an affective, friendly stance (e.g., smiling). When comprehension problems arose during the test I repeated the test questions and only clarified by glossing or paraphrasing when asked specifically, or using scripted clarification questions. Thus, the interaction was largely structured and presentational (e.g., non-interactive and one-directional) rather than interpersonal (e.g., bi-directional, interactive) (American Council on the Teaching of Foreign Languages,

2012), as I did not engage in any discussions with the test-taker. All speaking test materials are available at Burton (2021).

The present analysis is of question-response sequences from four test-takers on one question. For comparability, I selected the question on the test that first covered an abstract topic unfamiliar to the participants (question number seven of ten), which was: *What do young people in your country aspire for in their lives?* This question is approximately of B2 complexity on the CEFR according to an analysis of topic, function, and lexis (North et al., 2010). The target lexical unit *aspire* is a lower frequency, formal word listed as K4 in the BNC-COCA 1-25k corpus (Nation, 2016). This question was selected for analysis because its appearance in the test represented a shift in complexity from a concrete/near/familiar topic to one which was abstract/remote/unfamiliar (Skehan, 1998), and four comparable moments of breakdown and repair followed this particular test question.

Participants

I recruited the participants through an international institution in China. The four participants selected for this analysis all experienced comprehension problems and initiated repair sequences with the examiner on this particular test question. They were all Chinese females living in the same large city in northern China. Their ages ranged from 23 to 30, and each had been exposed to English first through primary or secondary education. I assigned the four participants pseudonyms as outlined in Table 1.

Table 1. Background and questionnaire data

Participant	Age	Grade of Exposure to English
Angela	30	5th
Beatriz	27	3rd
Carolina	27	7th
Deana	23	2nd

Data analysis

The four participants' video and audio files were analyzed using modified multimodal CA. Traditional strands of conversation analytic research encourage *unmotivated looking* (Heritage, 1988), or "an examination not prompted by prespecified analytic goals (not even that it be the characterization of an action), but by 'noticings' of initially unremarkable features of the talk or of other conduct" (Schegloff, 1996, p.172). The

impetus of this study, however, was to describe one specific type of social action—other-initiated repair—in one particular social context—the oral proficiency interview. That is to say, this study was indeed motivated by a particular research question, but the focus remained well within the CA framework as it was descriptive and not explanatory or experimental (Heritage, 1988). I transcribed all question-answer sequences in ELAN (Max Planck Institute, 2020). ELAN uses a tier-based annotation system which allows multiple parallel layers of information to be transcribed for each video. These tiers support annotations of individual segments which can represent words or actions. For this study, the salience of nonverbal features arose through repeated viewing of the recordings. I then devised tiers which best represented these features, allowing multiple behaviors to be annotated simultaneously. This resulted in 11 tiers which I used to annotate each video. A sample of the ELAN interface with annotations is given in Figure 1, and the 11 tiers are described in Table 2.

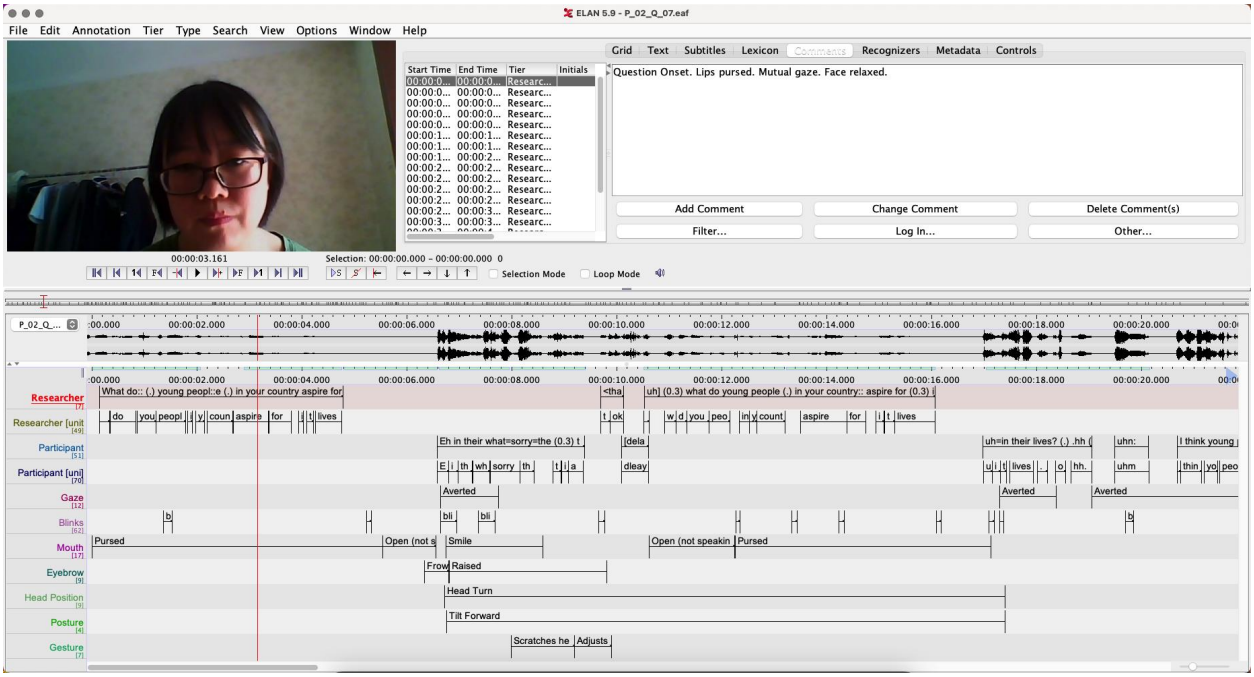

Figure 1. An example of the ELAN interface

Table 2. Tier descriptions

Tier	Label	Description
1-2	*Conversation Analysis*	Speech of examiner and test-taker were separately transcribed into two tiers using Jeffersonian CA transcription. In most cases these represented full TCUs.
3-4	*Words*	Speech as likewise separately transcribed word by word using orthographic transcription. Word boundaries were segmented by use of the waveform and a reduced speed audio recording.
5	*Gaze*	Gaze was annotated as averted from the moment of the first shift in eye direction until gaze was reconnected with the examiner.
6	*Blinks*	Blinks were segmented separately from gaze. Blink segments were annotated from the first moment the participant's eyelid began to fall and ended when the pupil again became visible.
7	*Mouth*	Three mouth behaviors were annotated. *Pursed Lips* were annotated when the participants' mouth was tightly closed, generally with the cheek muscles tight on each side of the lips. *Smiling* was annotated without distinguishing Duchenne smiles (contractions of both the zygomatic major muscle and the orbicularis oculi muscles; "smiling with the eyes") and non-Duchenne types (contracting only the zygomatic major muscle around the mouth). *Open (non-speaking)* was a category that appeared between speech segments where the participant held her mouth open without speaking.
8	*Eyebrow*	Two eyebrow movements were annotated. *Furrowed* brows were contracted, often with visible skin folds between the eyebrows. *Raised* indicated eyebrows lifted vertically away from the eyes.
9	*Head Position*	Head movements were classified as *head turn, head raise,* and *head lower*.
10	*Posture*	Posture referred broadly to the relationship between the participants' body and the camera. *Tilt Forward* occurred when the participant leaned from the neutral position towards the camera. *Tilt Backward* referred to movement away from the camera from the neutral position.
11	*Gesture*	Gestures were uncommon and segmented broadly as head nods or other general movements with the hand. *Head Nods* were annotated when the head rocked back and forth as a form of confirmation or assent. Other slight head movements that could not be confirmed to be nods were not annotated. *Iconic gestures* are gestures occurring with speech with a non-emblematic referential meaning (Kendon, 2004) and were annotated by describing them as closely as possible. These only occurred twice. Finally, *self-adaptors* are movements, generally of the hands, which may not co-occur with speech and generally are not representational in meaning (Ekman & Friesen, 1969). These were annotated as a description of the action taking place (e.g., scratches head).

For each behavior, the segment was annotated from the moment of the first muscle contraction or motion and held until that action relaxed or the body returned to a neutral position. For example, for eyebrows, a furrowed brow was annotated from the moment the eyebrows began to contract, and the segment ended when the eyebrows returned to a relaxed position. For posture, tilting forward was annotated from the moment the participants' body began to come closer to the camera, and the segment closed once the body returned to the neutral position. These extended segments rather than discrete movements were annotated in order to capture the presence of holds. Holds, which refer to behavior held in place for a period of time, may be singular gestures or combinations of behaviors that remain static (Floyd et al., 2016). This operationalization is slightly different from that of McNeill (1992) in that it is broader than the architecture of manual gesture. Because holds may be associated with one sustained behavior or multiple at once, these were not coded separately in the transcription, but are rather identified in the transcriptions and presented visually in Figures 2a and 2b.

Prior to annotating all of the data, one colleague experienced with CA transcription independently reviewed the verbal data and the segmented behaviors for 25% of the dataset. The colleague agreed on the segmentation scheme, the accuracy of the annotations, and the interpretability of the transcriptions. However, due to time and budget constraints, annotations were not conducted independently, and therefore calculations of intercoder reliability were not possible. Following this I then transcribed and annotated the remaining files. These files and transcriptions are available to view at Burton (2021).

The first analysis considered the microlevel features that appeared during repair sequences by using CA and the output from ELAN. I analyzed the annotation tiers for patterns of behavioral occurrences during trouble-sources, repair initiations, repair, and resolutions. By comparing participants' behaviors at these points, certain patterns emerged. I present excerpts of the findings using the time-aligned interlinear text transcription scheme available in ELAN. This scheme is slightly different from multimodal CA conventions to date, as it prioritizes the appearance of language and behaviors as time unfolds on the same scale. Thus, the transcript is temporally aligned, with horizontal unit measurements representing equal units of time (e.g., 1 cm measured horizontally anywhere in the transcript represents approximately 1 second).

In the transcriptions, I provide CA notation at the top with word-by-word timing in the line below to indicate the exact moment each individual word appeared for comparison with NVBs. The behavioral tiers correspond only to the test-taker. The appearance of each word or behavior is listed in brackets, and the extended use of the behavior is indicated using continuous lines. The end of the behavior is marked with a closing bracket, and the

extension of a behavior to subsequent lines is given with an arrow. When behaviors have continued from a previous section, these begin with a line without an opening bracket. To illustrate in the below example, Angela beings with her lips already smiling, marked using an open line, which continues throughout the question to the word 'country.' At this point the smile stops, shown using a closing bracket, and changes to pursed lips, indicated using opening brackets. At this point a frown also beings with an open bracket, and her body posture shifts shortly thereafter, as noted with an open bracket as she leans forward. The test taker blinks during the researcher's question five times with the examiner's words 'young,' 'people,' 'aspire,' 'for,' and at the conclusion of the question.

Example: Multimodal notation

```
Researcher     |_____________________________________________________________What do young people in your country (.) aspire for in their lives.]
Researcher [u..|____What]____do]____young] [peo..]() [.]  [_____________country]        [_______aspire]  [____________for]      [.] [th..][_______lives]
Participant    |
Participant [..|
Gaze           |
Blinks         |                    [.]            []                                        [__blink]     [bl..]                                  [_blink]
Mouth          |_____________________________________________________Smile] [_____________________________________________________________________________ Pursed→
Eyebrow        |                                                        [_____________________________________________________________________________ Frown→
Head Position  |
Posture        |                                                                  [_____________________________________________________ Tilt Forward→
Gesture        |
```

The second analysis considered individuals' differences in behavior during question-repair sequences. I adopted a descriptive quantitative approach for this analysis. Along with the detailed transcriptions, ELAN also provides count and duration information for each tier of information. I analyzed time sequences from the moment the examiner began asking the test question until the test-taker started their full answer without further repair initiations. Two numerical values were extracted from ELAN. Annotation duration percentages represented the ratio of time the test-taker spent performing the behavior relative to the total time measured. Annotation counts represented the number of times a participant performed the behavior in question. Because each participant's test question and repair sequence ranged from 19.86 to 47.76 seconds, I first scaled the behavioral counts by dividing by total time and multiplied this by 100. Scaling made values comparable between both participants and behaviors. Using this notation, 0 represents no use of the behavior whatsoever, and 100 represents one instance of the behavior per second.

Results

Broadly speaking, there were three key findings from the data analysis. The first two relate to the first research question, and the third relates to the second research question.

Finding 1: NVB during repair

The participants produced a wide range of behaviors during repair sequences, and the deployment of these was quite complex. Figures 2a and 2b show this information visually in annotation density graphs. These graphs show segments of time that were annotated for a particular behavior or observation. Researcher and Participant lines indicate sections of continuous, unbroken spoken utterances. Behavioral lines (e.g., gaze, mouth) indicate sustained, held behaviors (e.g., averted gaze, pursed lips). The bold red line indicates when resolution was achieved and the test-taker began answering the test question in full. Each tier indicates when utterances or behaviors appeared in a sequence timewise, as any vertical cross section represents the same moment of time for each tier. These graphs then allow a broad holistic view of the interactions between spoken language and behaviors, and thus can be seen as a type of 'partiture' of multimodal repair. These figures show that behaviors appeared and were released at different rates and in different combinations. Yet, as will be seen, some of these appeared to demonstrate identifiable patterns.

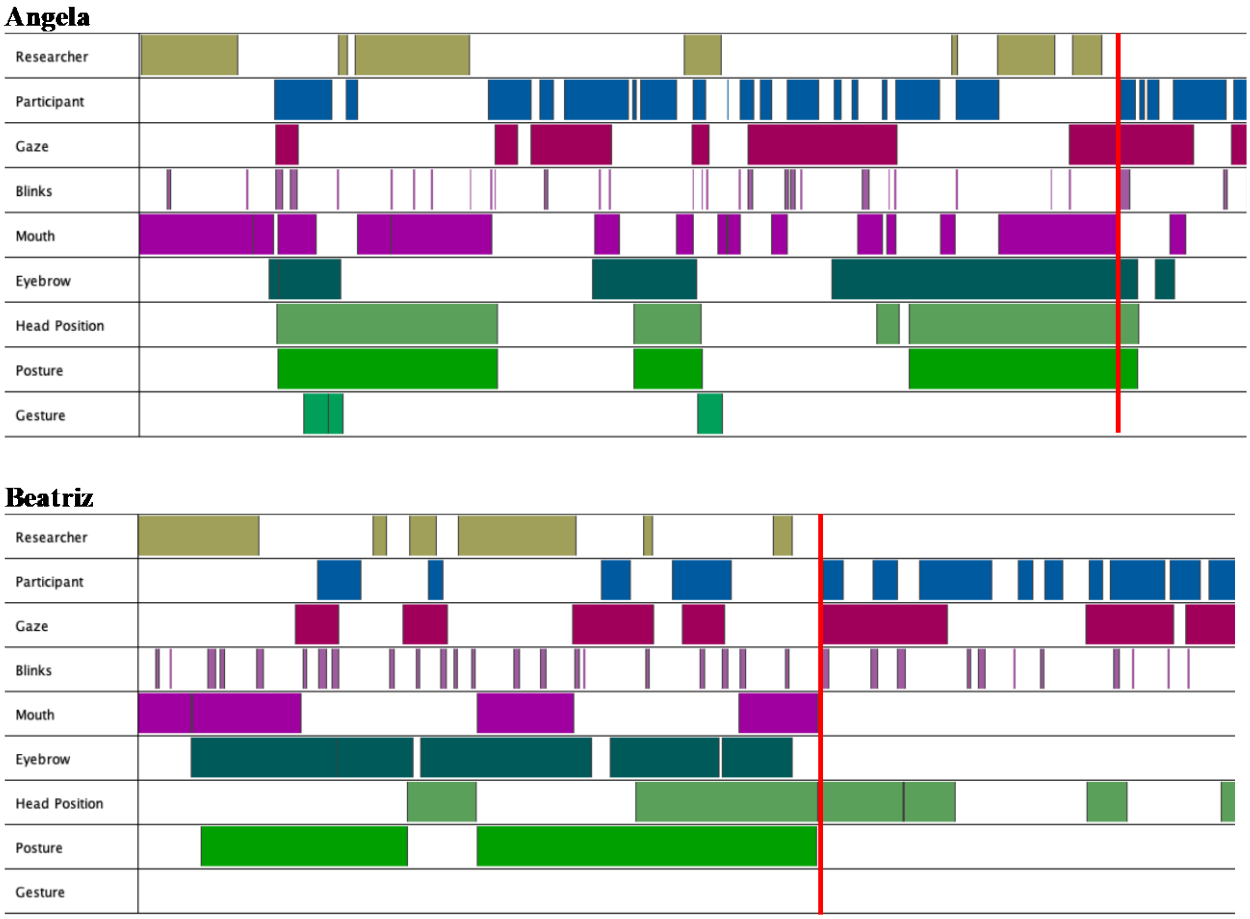

Figure 2a. Interactional sequences of Angela and Beatriz. Red lines indicate the beginning of test-takers' responses, non-repair TCUs.

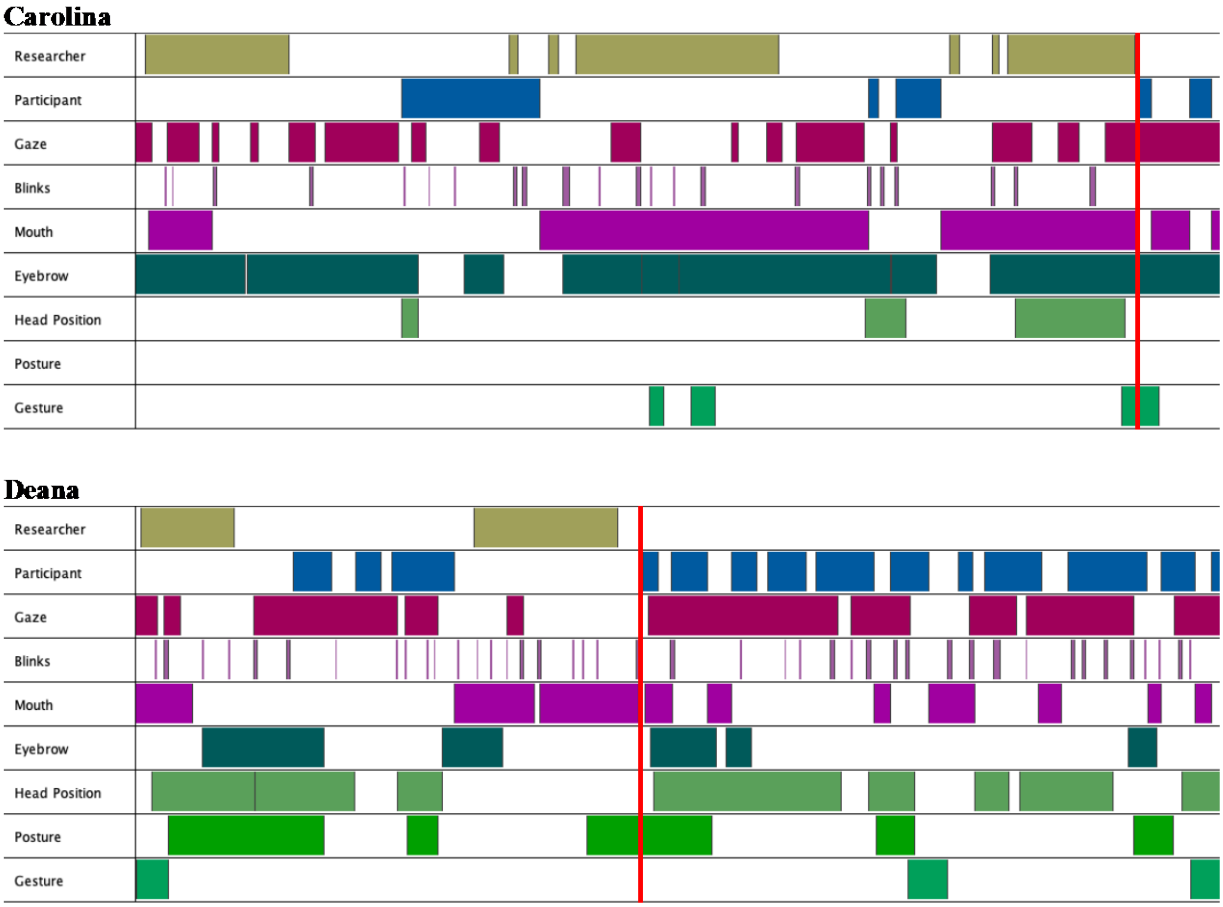

Figure 2b. Interactional sequences of Carolina and Deana. Red lines indicate the beginning of test-takers' responses, non-repair TCUs.

Angela's first repair initiation shows a clear example of a complex manifestation of behaviors arising after the trouble-source, as seen in Excerpt 1. Here Angela's behavior did not immediately change upon hearing the target word *aspire*, but rather shifted suddenly after 1.8 seconds of silence following the end of the question. This may be due to a delayed realization of the breakdown, or problems with the audio. Indeed, after the silence, she apologized, which is sometimes associated with problems hearing rather than problems understanding as she does not identify the trouble-source ('open class repair initiators;' Oloff, 2018). Nonetheless, just after the silence she completely shifted her behavior as she began her repair initiation by averting her gaze, blinking, opening her mouth, furrowing her brow, turning, and tilting toward the camera. Her gaze and eyebrow behavior are consistent with self-repair sequences documented in Tominaga (2013).

Excerpt 1–Angela–First Repair Initiation

```
Researcher      |    [_____________________________________________________What do:: (.) young peopl::e (.) in your country aspire for (.) in their lives.→
Researcher [u..|    [_What]_[_____________do]_[_____young]_[_____________people]_[]   [_your]_[_________country]_[_______________aspire]_[_____________for]    [_.]  [_t..]  [__lives→
Participant     |
Participant [..|
Gaze            |
Blinks          |                                        [_bl..]
Mouth           |_____________________________________________________________________________________________________________________Pursed→
Eyebrow         |
Head Position   |
Posture         |
Gesture         |

Researcher      |What d..]
Researcher [u..|__lives]
Participant     |                                                    [_____________________________________Eh in their what=sorry=the (0.3) this is a::→
Participant [..|                                                    [____Eh]_[__in]_[__their]_[_____what]_[_____________sorry]_[____the]            [_t..]_[.→
Gaze            |                                                      [_____________________________Averted]
Blinks          |                        [_.]                          [____blink]              [____blink]
Mouth           |_____________________________Pursed]_[_____________Open (not speaking)]      [_____________________________________________Smile]
Eyebrow         |                                                  [_________Frown]_[_____________________________________________________________Raised→
Head Position   |                                                                 [_____________________________________________________________Head Turn→
Posture         |                                                                  [____________________________________________________Tilt Forward→
Gesture         |                                                                                                  [_____________________Scratches head→
```

After initiating a repair, it is interesting to note that Angela used a self-adaptor, first scratching her head and then adjusting her glasses, as seen in Figure 3. Although this type of behavior is not overtly communicative and generally is thought to be associated with coping mechanisms (Ekman & Friesen, 1969), other explicitly communicative types of gestures—in particular referential—occurring during disfluencies often have pragmatic functions (Graziano & Gullberg, 2018). The onset of this motion from the word 'sorry' may thus have a pragmatic or emphatic function, albeit unconscious, that highlights Angela's ongoing comprehension difficulty.

Figure 3. Angela using two self-adapting gestures during repair initiation.

Not all test-takers exhibited such a complex and sudden onset of behaviors during these sequences. Carolina, for example, upon hearing the lexical unit *aspire for*, averted her gaze

slightly and shifted from a raised brow to a furrowed brow, but otherwise kept her posture the same. This eyebrow shift appears to be consistent with the flagging of trouble sources by paired test takers in Hırçın Çoban and Sert (2020). This visual response is shown in Figure 4. Her furrowed brows continued throughout the turn and the 2.94 seconds of silence before she responded with a repair initiation, which was preceded by her averted gaze. Her behavior across the sequence is given in Excerpt 2.

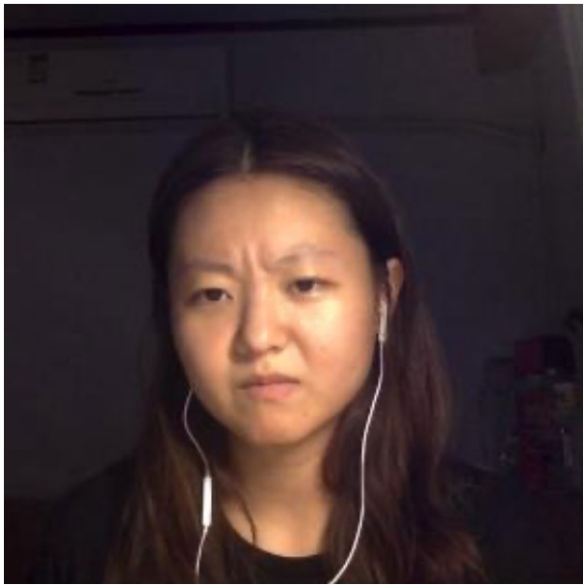

Figure 4. Carolina averting her gaze slightly and furrowing her brow after hearing the trouble-source.

Excerpt 2–Carolina–Trouble-Source

```
Researcher     |        [_____________________________________________ What do: (.) young people in your country aspire for (.) in their lives.]
Researcher [u..|        [__What][_______do][___young][__people][][__your][_________country][__________aspire][_____for]       [] [their][_____lives]
Participant    |
Participant [..|
Gaze           |_______Averted]            [_____________________Averted]           [Av..]                          [Ave..]                          [_________Averted→
Blinks         |                         []      []                                 [.]
Mouth          |           [_____________________________________________ Pursed]
Eyebrow        |_____________________________________________________________________________ Raised] [_____________________________________________ Frown→
Head Position  |
Posture        |
Gesture        |
_____________________________________________________________________________________________________________________________________

Researcher     |
Researcher [u..|
Participant    |                                                                                   [um (.) I'm- (.) I'm sorry (.) I didn't hear clearly (.) could yo..→
Participant [..|                                                                                   [____um]    []    [.][_______sorry]     [] [_didn't][___hear][cle..→
Gaze           |Aver..]        [_____________________________________________ Averted]          [___Averted]
Blinks         |  []                                                                                 []                  []                    []
Mouth          |
Eyebrow        |_____________________________________________________________________________ Frown]                                          [_____Raised→
Head Position  |                                                                                 [____Head Turn]
Posture        |
Gesture        |
```

Beatriz's behavior also changed, but after the onset of the word *country* instead of *aspire for*. As seen in Excerpt 3, trouble due to comprehension is clear from her repair initiation, as she asked for a clarification rather than an open-class initiator. Her behavior shifted near the end of the vocalization of the trouble-source, with her smile changing to a pursed

lip, brow furrowing, and a tilt forward. This appeared to signal her realization of a comprehension difficulty before this was verbalized. This behavioral response is shown in Figure 5.

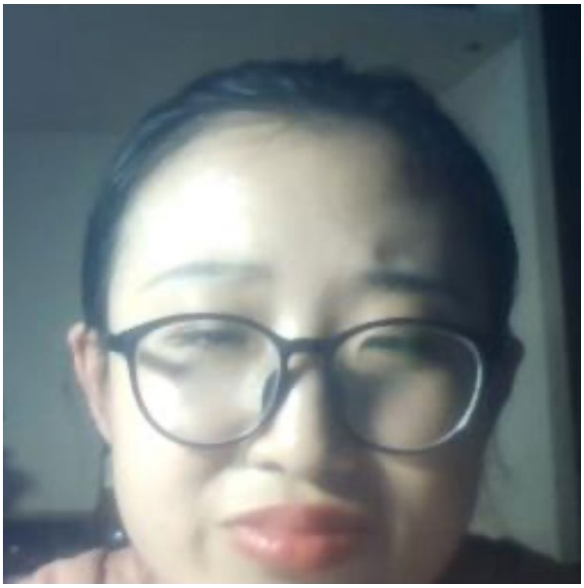

Figure 5. Beatriz's furrowed brow, pursed lips, and lean forward after the onset of the trouble-source.

Excerpt 3–Beatriz–First Repair Initiation

```
Researcher      |[_____________________________________________________What do young people in your country (.) aspire for in their lives.]
Researcher [u..]|[____What]([___do] [___young] [people]([)[.]   [___________country]       [_______aspire]  [___________for]      [_in][t..] [_______lives]
Participant     |
Participant [..]|
Gaze            |
Blinks          |                    [.]            []                                        [ blink]      [b..]                                  [blink]
Mouth           |[_____________________________________________Smile] [_____________________________________________________________________________Pursed→
Eyebrow         |                                                    [_____________________________________________________________________________Frown→
Head Position   |
Posture         |                                                        [_____________________________________________________________Tilt Forward→
Gesture         |


Researcher      |                                                                                                                  [_________Mhmm]
Researcher [u..]|                                                                                                                  [_________Mhmm]
Participant     |                                                  [_________________U::h (.) That's ambitions () yeah.]
Participant [..]|                                                   [__Uhh][.]     [___________ambitions]       [_yeah]
Gaze            |                           [                                      Averted]
Blinks          |                                   [b..]            [___blink]     [_blink]
Mouth           |_____________________________Pursed]
Eyebrow         |_____________________________________________________Frown][_____________________________________________Raised→
Head Position   |
Posture         |_____________________________________________________________________________________________Tilt Forward→
Gesture         |
```

In these excerpts, holds appeared quite regularly. The collection of features varied, but the participants' behavior was generally held from the trouble source until reaching some sort of resolution. This is illustrated in Figure 2a. In Beatriz's case, despite a backchannel from the examiner *mhmm*, the repair sequence was not resolved by the end of her TCU. Her embodied behavior did not change at this non-resolution; her raised brow and tilt forward continued. In fact, while other behaviors shifted throughout the repair sequence, her postural tilt forward continued until she reached final resolution. As opposed to the findings of Floyd et al. (2016), however, gaze was not always held in tandem with other facial expressions. Another clear example of the release of a hold due to a resolution is in

Excerpt 4. Here Angela reached consensus and released her mouth, brow, and head hold just after the first verbalization *uh::m* of her response to the test question.

Excerpt 4–Angela–Resolution

```
Researcher     |young people (0.2) in Ch..]
Researcher [u..|]   [ in][           China]
Participant    |                                                  [                uh::m.]       [  .hhh]     [     I think it-]                        [they..→
Participant [..|                                                  [                    Umm]       [  .hhh]     [  I ][ think][.]                        [  they]
Gaze           |_____________________________________________________________________________________________________________________________________Averted→
Blinks         |                                                [___________blink]
Mouth          |_____________________________________Open (not speaking)]                                                                                  [_____Smile→
Eyebrow        |_____________________________________________________________________Frown]                                  [_____________________Frown]
Head Position  |_____________________________________________________________________Head Turn]
Posture        |                                                                  Tilt Forward]
Gesture        |

Researcher     |
Researcher [u..|
Participant    |_____________________________________have many different opinions:: (.) and-]
Participant [..|___have][ many][__different][_____________________opinions][__________and]
Gaze           |_________________Averted]
Blinks         |                                                                        [.]
Mouth          |      Smile]
Eyebrow        |
Head Position  |
Posture        |
Gesture        |
```

Finally, though a full treatment is beyond the scope of this paper, there were some trends in the simultaneity of behaviors and speech. Apart from head nods, hand movements (both self-adaptors and iconic gestures) always co-occurred with the test-takers' speech, though in Angela's case, her self-adaptors occurred when initiating repair. This supports ongoing research which suggests that gestures generally occur during fluent speech in all speakers, though L2 speakers may gesture somewhat more than L1 speakers during disfluencies (Graziano & Gullberg, 2018). It is also interesting to note that except in the case of Carolina, gaze aversion almost exclusively coincided with when the test-taker was speaking, a phenomenon which is well documented in the normative gaze studies of Goodwin (1980). When listening to the examiner, the test-takers almost always maintained mutual gaze. Other facial movements, nevertheless, co-occurred with both speech and silence.

Finding 2: Variable release in holds

In moments of comprehension difficulty, holds often persist from the moment of comprehension difficulty until the resolution of the problem (Oloff, 2018; Seo & Koshik, 2010). A second finding in this study is that holds may be released intermittently and in various stages rather than continuously through the sequence. This was the case with Beatriz in Excerpt 5, whose brow release occurred before the release of her other behaviors. She held her tilt forward throughout the entire repair sequence and released her hold once she was ready to begin speaking. In this excerpt, Beatriz appeared to signal her final comprehension of the passage when she averted her gaze, blinked, and said *Ye::ahhh*. Just before this her hold ended as her mouth relaxed, her body posture returned

to a neutral position, and her head raised back. Interestingly, her brow behavior just preceded the end of the hold when her raised brow relaxed immediately with a blink after hearing the examiner's confirmation of her interpretation of the question.

Excerpt 5–Beatriz–Resolution

```
Researcher     |  [___ Yeah (0.3) Yeah]
Researcher [u..|  [_Yeah]         [y..]
Participant    |                                                  [_____________Ye::ahhh]                              [_________________Maybe? hh.]
Participant [..|                                                  [_________________Yeah]
Gaze           |                                              [_____________________________________________________________________________________Averted→
Blinks         |              [b..]                            [   blink]                                          [ blink]                        [blink→
Mouth          |_____________________________________________Pursed]
Eyebrow        |_______________Raised]
Head Position  |_____________________________________Head Turn][_____________________________________________________________________________Head Raise]
Posture        |_________________________________Tilt Forward]
Gesture        |


Researcher     |
Researcher [u..|
Participant    |                 [_____________________________________a different way- (.) age have different]
Participant [..|
Gaze           |                                      Averted]
Blinks         |]                                                          [b..]         [_blink]
Mouth          |
Eyebrow        |
Head Position  |[_____________________________________________Head Lower]
Posture        |
Gesture        |
```

This variable release pattern was even more evident in Carolina's resolution, seen in Excerpt 6. Here, as soon as she heard the clarification *want to do*, which replaced *aspire for*, she blinked and then averted her gaze, which may have signaled her acceptance of the upcoming turn (Goodwin, 1980). As she began her turn there was a near simultaneous relaxing of her mouth, head, and a brief head nod, which appeared to show her comprehension. Her brow, however, remained furrowed for an additional 3.2 seconds after she had indicated comprehension rather than being released, as in the case of Beatriz.

Excerpt 6–Carolina–Resolution

```
Researcher      |   [_____________________________________What do young people  want to DO. (0.2) In the future.]
Researcher [u..|   [ What][__do_][young][peo..][][.] [_____country][_______want][__to][________do]       [][__the][_________future]
Participant     |                                                                                                                  [_____Yeah.]
Participant [..|                                                                                                                  [_____Yeah]
Gaze            |_____________Averted]                        [________Averted]                    [_____________________________________________________Averted→
Blinks          |         [.]                                                                  [b..]
Mouth           |                                                                                          Open (not speaking)]          [Open (not speaking)→
Eyebrow         |_____________________________________________________________________________________________________________________________________Frown→
Head Position   |          [_____________________________________________________________Head Turn]
Posture         |
Gesture         |                                                                                                  [_____________________Head Nod]
_____________________________________________________________________________________________________________________________________________________

Researcher      |
Researcher [u..|
Participant     |            [_____________Umm]                                                                          [___hhh.]
Participant [..|            [_____________Umm]                                                                          [(laug..]
Gaze            |_____________________________________________________________________________________________________________________________Averted
Blinks          |                                                                                                  [____blink]          [.]                [.]
Mouth           |Open (not..]                  [                            Pursed]                                  [                                      Smile→
Eyebrow         |_____________________________________________________Frown]
Head Position   |                                                                                                                   [_____________Head Raise→
Posture         |                                                                                                         [_____________________________→
Gesture         |
_____________________________________________________________________________________________________________________________________________________

Researcher      |
Researcher [u..|
Participant     |    [_________________uhh= ↑I'm not sure.]
Participant [..|   [_______Uhh][I'm][___not][_______sure]
Gaze            |                            [_______Averted→
Blinks          |             []           [.]
Mouth           |                                   Smile→
Eyebrow         |
Head Position   |.]
Posture         |_____________________________________→
Gesture         |
_____________________________________________________________________________________________________________________________________________________
```

The variable release of a hold was perhaps most evident in Deana's first repair initiation, seen in Excerpt 7. Here Deana began her hold at the onset of the test question, yet the onset of the hold was not simultaneous. Her hold began as a lowering of the head, then tilt towards the camera, and finally a furrowing of her brow accompanied with a blink at the trouble-source. This brow/head/posture hold was not held until resolution but rather ended just after she initiated a repair sequence. In this case, her head turn relaxed shortly after her brow and posture.

Excerpt 7–Deana–Repair initiation

```
Researcher     |        [_____________________________________________ What do: young people in your country (.) aspire for (.) in their lives.]
Researcher [u..|       [_What]  [___do]  [__young]  [_________people]  []  [y..]  [_____country]  [_____________aspire]  [_____for]  [_in]  [t..]  [_________lives]
Participant    |
Participant [..|
Gaze           |                      Averted]        [            Averted]
Blinks         |                          []          [b..]                                                  []                                  []
Mouth          |_____________________________________________________________________Pursed]
Eyebrow        |                                                                                             [_____________________________________________________Frown→
Head Position  |                     [_____________________________________________________________________________________________________________________Head Lower→
Posture        |                                   [_____________________________________________________________________________________________________Tilt Forward→
Gesture        | [_________________________________Head Nod]
_______________________________________________________________________________________________________________________________________________________________________

Researcher     |
Researcher [u..|
Participant    |                                                   [                        U(h)m (0.3) Like what-]                                  [about the..→
Participant [..|                                                   [_____________Uhm]       [_________like]  [_____what]                             [_____about→
Gaze           |     [_____________________________________________________________________________________________________________________________________Averted→
Blinks         |     [.]                                  [.]                                                                  []
Mouth          |
Eyebrow        |_____________________________________________________________________________Frown]
Head Position  |Hea..]  [_____________________________________________________________________________________________________________________Head Turn]
Posture        |_____________________________________________________________________Tilt Forward]
Gesture        |


Researcher     |
Researcher [u..|
Participant    |_____________young-]              [_____________________________ch- (0.2) the young generation are ↑chasing?]
Participant [..|a..]  [.]  [_____young]             [ch-]     [uhh]  [_the]  [_________young]  [_________generation]  [_are]  [_____________chasing]
Gaze           |_____________________________Averted]        [_____________________________Averted]
Blinks         |                                  []        [.]                        []       []
Mouth          |                                                                                                                                  [.→
Eyebrow        |                                                                                                                 [          Raised→
Head Position  |                                    [                                          Head Raise]
Posture        |                                                 [                          Tilt Back]
Gesture        |
_______________________________________________________________________________________________________________________________________________________________________
```

After initiating a repair her behavior was mostly relaxed, despite the fact that resolution had not been reached. She resumed the hold with an open mouth and re-established gaze at the end of her repair, and this mouth behavior was held throughout the following sequence in Excerpt 8. In the resolution sequence, she maintained eye contact with the examiner and held her mouth position until the clarification *want to do*, after which she formed a pursed mouth until the end of the sequence. Interestingly, near the end of the question *in their lives (.) in the future*, she tilted forward and held this pose until after she began her answer sequence. The rest of her behavior, nevertheless, shifted once she began speaking, with a blink followed by averted gaze, furrowed brow, and a turned head. This is shown in Figure 6. While her averted gaze likely indicated her acceptance of the turn, the furrowed brow and turned head, which to this point have been associated with comprehension problems, may have been due to her lexical search before she began her answer in earnest with *<maybe=it's a->*, as documented in the lexical searches in Tominaga (2013). Thus, defining exact hold boundaries may be complicated by other social or cognitive processes co-occurring with the resolution of repair. One of these social processes is the testing context itself. The asymmetrical, rigid question-repair-answer format may possibly interrupt holding patterns that normally would have been held in conversational, interpersonal discourse.

Excerpt 8–Deana–Resolution

```
Researcher      |  [_____________________________mmhmm (0.20) yeah. (.) what do young people in your country:::=want to do:: (0.6) u::h in their lives (.) in the future.→
Researcher [u..|  [___mmhmm]     [_____yeah]     [w..][.][__young][_people][] [your][_____country][___want][__to][_________do]                    [_____uhh][_in][→
Participant     |
Participant [..|
Gaze            |                                                  [___________Averted]
Blinks          |       []              []                    []              [.]                  []                                          [.]          []
Mouth           |_____________________________________________________________Open (not speaking)]      [_____________________________________________Pursed→
Eyebrow         |_____________________________Raised]
Head Position   |
Posture         |
Gesture         |
_____________________________________________________________________________________________________________________________________________________

Researcher      |mmhmm (0.20) yeah. (.) what do young..]
Researcher [u..|.][_______lives][][] [_________future]
Participant     |                                                                                [_________________mmmm]                [_________________ts .hhh well ↑ I think.→
Participant [..|                                                                                [_________________mmmm]                [] [_________.hhh][__well][___I][___think→
Gaze            |                                                                                      [_____________________________________________________Averted→
Blinks          |            [.]                                                  [b..]                                  [blink]
Mouth           |_____________________________________________________________Pursed]
Eyebrow         |                                                                                                [                                                Frown→
Head Position   |                                                                                                     [_____________________________________Head Turn→
Posture         |_____________________________________________________________________________________________________________________________Tilt Forward→
Gesture         |
_____________________________________________________________________________________________________________________________________________________

Researcher      |
Researcher [u..|
Participant     |ts ...]                                [_____u:h (0.2)_]             [_____________________________a little cliche, bu::t]
Participant [..|_think]                                 [____uhh]     [_____maybe][][___a]            [.][____little][_________________cliche][_____but]
Gaze            |_____________________________________________________________________________________________________________________________Averted→
Blinks          |                                                  []                                                    []                 []
Mouth           |      [____________Open (not speaking)]
Eyebrow         |____________Frown]             [____________________________Frown]
Head Position   |                                                                                                                                  Head Turn→
Posture         |Tilt Forward]
Gesture         |
_____________________________________________________________________________________________________________________________________________________
```

Figure 6. Deana's behavior after beginning her turn resolution sequence: averted gaze, furrowed brow, and turned head.

Finding 3: Evidence of idiosyncratic behavior

An analysis of behavioral data shows the participants varied quite substantially in almost all behaviors. Table 3 contains the annotation duration percentages and Table 4 contains

the scaled frequency of annotations for the four participants. Figure 7 displays these points in a spaghetti plot. Of the participants, Angela was perhaps the most overall expressive test-taker. Her head, eye, and mouth remained in an active state for nearly 25% of the time segment. Beatriz was also quite expressive in her eyebrow movements and her leaning posture, but otherwise maintained more mutual gaze and did not use any gestures. Carolina, on the other hand, showed the most expressive eye, eyebrow, and mouth behavior, averting her gaze nearly 20% of the time, with gaze aversions happening about once every two seconds. Her eyebrows remained furrowed or raised nearly 40% of the interview as well. Her head and posture remained mostly fixed, however, despite somewhat occasional nods. Deanna was the least expressive of the test-takers. Her behavior durations only lasted about 10% of the interview, though she was the most frequent blinker.

Table 3. Annotation duration percentages

Participant	Time	Averted Gaze	Blinks	Mouth	Eyebrow	Head	Posture	Gesture
Angela	47.76	16.21	3.91	26.32	22.02	24.69	23.60	2.98
Beatriz	23.51	10.36	5.00	16.19	27.26	12.25	26.08	0
Carolina	26.15	20.07	3.48	28.23	40.41	8.00	0	2.63
Deana	19.86	11.24	2.38	11.46	8.82	12.00	11.76	1.57
Mean	29.32	14.47	3.69	20.55	24.63	14.24	15.36	1.80

Table 4. Scaled frequency of annotations

Participant	Gaze	Blinks	Mouth	Eyebrow	Head	Posture	Gesture
Angela	12.56	60.72	29.31	8.38	8.38	6.28	6.28
Beatriz	21.27	97.83	17.01	21.27	12.76	8.51	0
Carolina	61.19	84.13	11.47	30.59	11.47	0	11.47
Deana	20.14	110.78	15.11	10.07	15.11	15.11	5.04
Mean	28.79	88.37	18.23	17.58	11.93	7.48	5.70

Figure 7. Profiles of individuals' NVBs.

Discussion

In L2-speaking tests, a major feature of the interview structure is the appearance of comprehension breakdowns and resulting repair sequences. While there has been research on the verbal response structure of repair sequences in OPIs (Kasper & Ross, 2007; Kim & Suh, 1998; Kim & Park, 2015; Roever & Kasper, 2018), with this study I expanded the focus to investigate the NVBs that arose in test-taker-initiated repair sequences and comprehension resolution sequences, thus adding to the current body of research (e.g., Jenkins & Parra, 2003; Tominaga, 2013) and ongoing calls for a broader focus in this area (Plough, 2021; Plough et al., 2018; Roever & Kasper, 2018). This has thus offered a glimpse of how L2 speakers may communicate their problematic understanding

(explicitly or implicitly) using a complex constellation of behaviors, and how these individuals may vary.

Many of the behaviors that appeared in this study have been previously documented in non-testing settings. Furrowed brows (Faraco & Kida, 2008; Oloff, 2018) and head turning and changes in posture (Seo & Koshik, 2010; McDonough et al., 2019; Oloff, 2018) have been documented during comprehension difficulties with L2 learners. In these varying contexts, these behaviors often manifested as holds (Floyd et al., 2016). There were also some indications of the co-occurrence of speech and gesture, as well as speech and averted gaze (Graziano & Gullberg, 2018; Goodwin, 1980). Some forms, such as the 'head poke' (Seo & Koshik, 2010) or extended gaze holds (Floyd et al., 2016) were nonetheless not present in these samples. Thus, a question that emerged is whether the interactional rigidity of the testing context affected the types of NVBs test-takers exhibited. It is possible that the necessity for test-takers to initiate repair in a more structured interview-style format reduced or limited the opportunities for participants to use certain groupings of nonverbal resources. More research comparing behaviors arising in different discourse styles is necessary to reveal key differences in these contexts.

Some of the behaviors documented in this study have received less attention in past research. For one, although McDonough et al. (2019) found no major differences in mouth behavior between episodes of understanding and nonunderstanding, the test-takers in this study often pursed their lips or slightly opened their mouths during hold sequences which began following trouble-sources. Likewise, there was some indication that blinks co-occurred with problematic utterances, but their sheer ubiquity complicates attempts at extrapolation. Finally, Angela was the only test-taker to use self-adaptors (Ekman & Friesen, 1969) during repair requests. By brushing away her hair and adjusting her glasses simultaneously with brow changes, head turns, and tilting forward during repair, she may have further emphasized her ongoing problems in communication. Finally, although head nods have also been found to coincide with some moments of nonunderstanding (McDonough et al., 2019), in this study they were largely absent except when Carolina was backchanneling.

Another important finding in this study is that holds may be more complex than previously reported in the literature. To date, Floyd et al. (2016), Oloff (2019), and McDonough et al. (2019) have reported the presence of holds during comprehension breakdowns that last from the beginning of a comprehension problem until its resolution. Figures 2a and 2b show that this was partially the case in this dataset as well, with behaviors generally beginning at trouble sources and ending at the beginning of response sequences. However, these holds were sometimes released during or after repair initiation, and the final release was generally not immediate for all behaviors (e.g., Angela)

but sometimes staggered, with some behaviors being released after others (e.g., Carolina and Deana). For this reason, care should be taken if holds are used to indicate comprehension problems. It is possible that in the context of L2 speaking tests, response sequences marked by disfluencies such as repair may muddy the boundaries of hold sequences because test-takers could very well desire to feign understanding to the examiner even when this may not be the case.

The results also show that although speakers tend to use a similar range of behaviors, the duration, frequency, and patterns in which they occur may be highly individualized. However, without baseline behavioral indices, it is impossible to ascribe any causal relationships between these metrics and communication difficulties. One may speculate that some degree of expressiveness may be due to the difficulty of the test question. In the post-test questionnaire, Angela, Beatriz, and Carolina all reported that this question was difficult and that their understanding was low, while Deana reported that the test question was of neutral difficulty and her understanding was good. These three test-takers were markedly more expressive than Deana. Although Deana blinked more frequently, her behavior was more relaxed and stable across the test. This type of expressiveness may then be associated with an individual's ongoing struggle to re-establish intersubjectivity with the examiner. Nevertheless, these data show substantially different constellations of behavior while these test-takers navigated comprehension of these test questions. The trends here indicate the importance of considering individual variability in behavior when generalizing to populations, as each person demonstrated somewhat different patterns when dealing with repair.

Conclusion

This paper has important implications for the assessment of second language speech. For one, the behavioral descriptions in this paper show that NVB does not exhibit a one-to-one correspondence with particular functions or meaning, and as such any integration of it into rating scales must be theoretically-informed and avoid its isolation and quantification (e.g., "test-taker uses few/some/many gestures") (cf. Jungheim, 2001). For example, the eyebrow furrowing as described in this study may be a precursor to an other-initiated repair sequence, or it could very well relate to a lexical search in self-repair or an affective stance. Likewise, nodding may be evidence of backchanneling, which can indicate comprehension, or instead it may possibly be an individual's idiosyncratic representation of beat gestures while on camera. As Plough (2021) emphasized, the underlying use of NVB is context-dependent and fluid. If rating scales incorporate NVB, descriptors should be written in a way such that the function is described and allowed to be reflected through multiple channels. An example may be "The speaker uses

appropriate verbal and nonverbal cues to indicate comprehension problems and to initiate clarification sequences." Words such as "appropriate," though somewhat subjective, can be clarified to raters in rater training through the use of benchmarked samples that reflect varying patterns in behavior that aid or inhibit communicative effectiveness. These benchmarked samples can be chosen and detailed through methods such as the multimodal conversation analysis described in this paper.

This study also has implications for the ongoing theoretical development of the construct of IC for language assessment. The most current conceptualization of this construct is the tree-shaped diagram in Galaczi and Taylor (2018, p.227). This diagram includes language rooted in speech acts, events, and situational contexts, and describes the functions of IC as branches including turn management, topic management, non-verbal behavior, breakdown repair, and interactive listening. However, as I have described in this study, non-verbal behavior alone cannot constitute its own branch, as it is never isolated from its context to the extent that behaviors can independently form functions of their own, nor can they independently reveal the agency or intention of the speaker systematically (Plough, 2021). Nonverbal behavior generally aligns with speech (Graziano & Gullberg, 2018) and synthesizes with it to produce meaning for the listener (Kita et al., 2007). As seen in examples here, nonverbal behavior jointly assists language in resolving breakdown and repair sequences rather than the verbal mode alone, and likely also combines with the other branches of the tree, such as the use of gaze to allocate turns during turn management (Nakatsuhara, 2011; Rossano, 2012). Thus, instead of assigning NVB to a separate branch on the tree, I envision each leaf on each branch as colored by both verbal and nonverbal channels, suggesting that each function can be satisfied by a combination of verbal and nonverbal communication.

This study is not without limitations. For one, working at such a micro-level of detail in any discourse context limits the number of cases that can be analyzed. A second limitation is that breakdowns and repair sequences detailed in this study may be category-bound (Sacks, 1984) to the role of 'being an examinee' in oral proficiency tests. Interviews, including oral proficiency tests, appear to be characterized by a particular set of institutional practices and routine discourse types that differ substantially from a more naturalistic classroom setting. This departure from naturalistic settings may have even more pronounced effects on comprehension due to the online format (Nakatsuhara, Inoue, Berry, & Galaczi, 2021). These tests are bound by interactional asymmetry induced by both the increasingly abstract task and the presence of a native speaking examiner (Young & Milanovic, 1992). This naturally leads to particular types of repair initiations and repair sequence patterns which may occur less frequently in the presence of face-to-face, 'amicable' interlocutors who may monitor and adjust their language as turns unfold.

This online institutionalized interaction may also lead to somewhat different nonverbal reactions during breakdown and repair.

Despite these limitations, this study has added evidence to show that audio transcripts alone may be insufficient when describing the turn-by-turn unfolding of repair sequences in online testing contexts. Each person utilized a different repertoire of NVBs to further indicate (and potentially communicate) a breakdown in comprehension. This visual overlay is a crucial part of communication and undoubtedly influences the outcomes of repair sequences. While CA work has always allowed the inclusion of nonverbal information in its transcription schemes (Sacks et al., 1974), current technology such as ELAN makes its analysis much more amenable to study. Ongoing work in this area will need to build on this type of research by building a better understanding of baseline NVBs—the idiosyncratic behaviors of each individual—in order to confirm which behaviors appear as a result of differing discourse patterns. On the same token, studies of learners at different proficiency levels (and, in particular, longitudinal studies) should be undertaken to describe nonverbal aspects of communicative ability along a developmental continuum. As opposed to verbal ability, nonverbal ability appears to be innate and can be used by learners at any level to support communication. How learners leverage these behaviors may change as they become more proficient in a second language. This type of analysis can then be used to inform scale construction for the rating of second language speech and the selection of appropriate samples for benchmarking.

Acknowledgements

I would like to thank Dr. India Plough and Dr. Koen van Gorp for providing motivation, guidance, and insight during this project. I would also like to thank the anonymous reviewers of this manuscript as well as the editors of PLTA for their kind recommendations and support.

Declaration of interest statement

No potential competing interest is reported by the author.

Open data and materials

The author has made materials and data available online in an Open Science Framework repository (Burton, 2021).

References

American Council on the Teaching of Foreign Languages. (2012). *ACTFL proficiency guidelines*. ACTFL. https://www.actfl.org/sites/default/files/guidelines/ACTFLProficiencyGuidelines2012.pdf

Benjamin, T., & Mazeland, H. (2012). Conversation analysis and other-initiated repair. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal1310

Burch, A. R., & Kley, K. (2020). Assessing interactional competence: The role of interusbjectivity in a paired-speaking task. *Papers in Language Testing and Assessment*, *9*(1), 25–63. http://www.altaanz.org/uploads/5/9/0/8/5908292/2020_9_1__2_burch_kley.pdf

Burton, J. D. (2021). *Question complexity, nonundersetanding, and nonverbal behavior* (OSF F2JXK, Version V1) [Data set]. Open Science Framework. https://doi.org/10.17605/OSF.IO/F2JXK

Cooke, S. (2015). *Configuring the game of speaking: Interactional competence in the IELTS Oral Proficiency Interview across two modes of response* [Unpublished master's dissertation]. Lancaster University.

Dingemanse, M., Kendrick, K. H., & Enfield, N. J. (2016). A coding scheme for other-initiated repair across languages. *Open Linguistics*, *2*(1), 35–46. https://doi.org/10.1515/opli-2016-0002

Drew, P. (1997). "Open" class repair initiators in response to sequential sources of troubles in conversation. *Journal of Pragmatics*, *28*(1), 69–101. https://doi.org/10.1016/s0378-2166(97)89759-7

Ducasse, A. M., & Brown, A. (2009). Assessing paired orals: Raters' orientation to interaction. *Language Testing*, *26*(3), 423–443. https://doi.org/10.1177/0265532209104669

Ekman, P., & Friesen, W. V. (1969). The repertoire of NVB: Categories, origins, usage, and coding. *Semiotica*, *1*(1), 49–98. https://doi.org/10.1515/9783110880021.57

Faraco, M., & Kida, T. (2008). Gesture and the negotiation of meaning in a second language classroom. In S. G. McCafferty & G. Stam (Eds.), *Gesture: Second language acquistion and classroom research* (pp. 280–297). Routledge. https://doi.org/10.4324/9780203866993

Floyd, S., Manrique, E., Rossi, G., & Torreira, F. (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

Galaczi, E., & Taylor, L. (2018). Interactional competence: Conceptualisations,

operationalisations, and outstanding questions. *Language Assessment Quarterly*, *15*(3), 219-236. https://doi.org/10.1080/15434303.2018.1453816

Gan, Z., & Davison, C. (2011). Gestural behavior in group oral assessment: A case study of higher- and lower-scoring students. *International Journal of Applied Linguistics, 21*(1), 95-120. https://doi.org/10.1111/j.1473-4192.2010.00264.x

Gardner, R., & Wagner, J. (Eds.). (2004). *Second language conversations*. Continuum.

Goodwin, C. (1980). Restarts, pauses, and the achievement of mutual gaze at turn-beginning. *Sociological Inquiry, 50*, 272–302. https://doi.org/10.1111/j.1475-682X.1980.tb00023.x

Graziano, M., & Gullberg, M. (2018). When speech stops, gesture stops: Evidence from developmental and crosslinguistic comparisons. *Frontiers in Psychology*, *9*(879). https://doi.org/10.3389/fpsyg.2018.00879

Greer, T., & Potter, H. (2008). Turn-taking practices in multi-party EFL oral proficiency tests. *Journal of Applied Linguistics*, *5*(3), 297–320. https://doi.org/10.1558/japl.v5i3.297

He, A. W., & Young, R. (1998). Language proficiency interviews: A discourse approach. In R. Young & A. W. He (Eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency* (pp. 1–24). Benjamins.

Hellermann, J. (2011). Members' methods, members' competencies: Evidence of language learning in longitudinal studies of other-initiated repair. In J. Hall, J. Hellermann, & S. Pekarek Doehler (Eds.), *The Development of Interactional Competence* (pp. 147–172). Multilingual Matters.

Heritage, J. (1988). Explanations as accounts: A conversation analytic perspective. In C. Antaki (Ed.), *Analyzing lay explanation: A case book of methods* (pp. 127–144). Sage.

Hırçın Çoban, M., & Sert, O. (2020). Resolving interactional troubles and maintaining progressivity in paired speaking assessment in an EFL context. *Papers in Language Testing and Assessment*, *9*(1), 64–94. http://www.altaanz.org/uploads/5/9/0/8/5908292/2020_9_1__3_hircin-soban_sert.pdf

Institute, Max Planck. (2020). *ELAN (Version 5.9)*. The Language Archive.

Jenkins, S., & Parra, I. (2003). Multiple layers of meaning in an Oral Proficiency Test: The complementary roles of nonverbal, paralinguistic, and verbal behaviors in assessment decisions. *The Modern Language Journal*, *87*(1), 90–107. https://doi.org/10.1111/1540-4781.00180

Jungheim, N. O. (2001). The unspoken element of communicative competence: Evaluating language learners' nonverbal behavior. In T. Hudson & J. Brown (Eds.), *A focus on language test development: Expanding the language proficiency construct*

across a variety of tests (pp. 1-35). University of Hawaii, Second Language Teaching and Curriculum Center.

Kappas, A., Krumhuber, E., & Küster, D. (2013). Facial behavior. In J. A. Hall & M. L. Knapp (Eds.), *Nonverbal communication* (pp. 131–165). De Gruyter.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press. https://doi.org/10.5860/choice.42-5687

Kim, K., & Suh, K. (1998). Confirmation sequences as interactional resources in Korean language proficiency interviews. In R. Young & A. W. He (Eds.), *Talking and testing. Discourse approaches to the assessment of oral proficiency* (pp. 297–353). John Benjamins.

Kim, S. H., & Park, I. (2015). Test taker-initiated repairs in an English oral proficiency exam for international teaching assistants. *Text and Talk*, *35*(2), 237–262. https://doi.org/10.1515/text-2014-0036

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, *22*(8), 1212–1236. https://doi.org/10.1080/01690960701461426

Lazaraton, A. (2002). *A qualitative approach to the validation of oral language tests*. Cambridge University Press.

May, L. (2011). Interactional competence in a paired speaking test: Features salient to raters. *Langauge Assessment Quarterly*, *8*(2), 127-145. https://doi.org/10.1080/15434303.2011.565845

McDonough, K., Trofimovich, P., Lu, L., & Abashidze, D. (2019). The occurrence and perception of listener visual cues during nonunderstanding episodes. *Studies in Second Language Acquisition*, *41*(5), 1151–1165. https://doi.org/10.1017/S0272263119000238

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

Mortensen, K. (2012). Visual initiations of repair—some preliminary observations. In K. Ikeda & A. Brandt (Eds.), *Kansai University International Symposium: Challenges and new directions in the micro-analysis of social interaction* (pp. 45–50). Kansai University.

Nakatsuhara, F. (2011). Effects of test-taker characteristics and the number of participants in group oral tests. *Language Testing*, *28*(4), 483–508. https://doi.org/10.1177/0265532211398110

Nakatsuhara, F., Inoue, C., Berry, V., & Galaczi, E. (2021). Video-conferencing speaking tests: Do they measure the same construct as face-to-face tests? *Assessment in Education: Principles, Policy, & Practice*. Advance online publication.

https://doi.org/10.1080/0969594X.2021.1951163

Nakatsuhara, F., Inoue, C., & Taylor, L. (2021). Comparing rating modes: Analyzing live, audio, and video ratings of IELTS speaking test performances. *Language Assessment Quarterly*, *18*(2), 83– 106. https://doi.org/ 10.1080/15434303.2020.1799222

Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. Amsterdam.

North, B., Ortega, A., & Sheehan, S. (2010). *A core inventory for general English*. British Council/European Association for Quality Language Services.

Oloff, F. (2018). "Sorry?"/"Como?"/"Was?" – Open class and embodied repair initiators in international workplace interactions. *Journal of Pragmatics*, *126*, 29–51. https://doi.org/10.1016/j.pragma.2017.11.002

Plough, I. (2021). A case for NVB: Implications for construct, performance, and assessment. In M. R. Salaberry & A. R. Burch (Eds.), *Assessing Speaking in Context: Expanding the Construct and its Applications* (pp. 50–69). Multilingual Matters.

Plough, I., Banerjee, J., & Iwashita, N. (2018). Interactional competence: Genie out of the bottle. *Language Testing*, *35*(3), 427-455. https://doi.org/10.1177/0265532218772325

Roever, C., & Dai, D. W. (2021). Reconceptualizing interactional competence for language testing. In M. R. Salaberry & A. R. Burch (Eds.), *Assessing Speaking in Context: Expanding the Construct and its Applications* (pp. 23–49). Multilingual Matters.

Roever, C., & Kasper, G. (2018). Speaking in turns and sequences: Interactional competence as a target construct in testing speaking. *Language Testing*, *35*(3), 331–355. https://doi.org/10.1177/0265532218758128

Ross, S. J. (2017). Interviewing for language proficiency: Interaction and interpretation. In *Interviewing for Language Proficiency: Interaction and Interpretation*. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-60528-9

Rossano, F. (2012). *Gaze behavior in face-to-face interaction* [Unpublished doctoral dissertation]. Radboud University. http://hdl.handle.net/2066/99151

Sacks, H. (1984). On doing "being ordinary." In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 413–429). Cambridge University Press. https://doi.org/10.1017/cbo9780511665868.024

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735. https://doi.org/10.2307/412243

Salaberry, R., & Burch, A. R. (Eds.) (2021). *Assessing speaking in context: Expanding the construct and its applications*. Multilingual Matters.

Salaberry, R., & Kunitz, S. (Eds.) (2019). *Teaching and testing L2 interactional competence:*

Bridging theory and practice. Routledge.

Schegloff, E. A. (2000). When "others" initiate repair. *Applied Linguistics*, *21*(2), 205–243. https://doi.org/10.1093/applin/21.2.205

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*(2), 361–382. https://doi.org/10.2307/413107

Seedhouse, P. (2013). Oral proficiency interviews as varieties of interaction. In S. J. Ross & G. Kasper (Eds.), *Assessing Second Language Pragmatics* (pp. 199–219). Palgrave Macmillan. https://doi.org/10.1057/9781137003522_8

Seedhouse, P., & Egbert, M. (2006). The interactional organisation of the IELTS Speaking Test. In P. McGovern & S. Walsh (Eds.), *IELTS Research Reports* (Vol. 6; pp. 1–45). IELTS Australia and British Council.

Seo, M. S., & Koshik, I. (2010). A conversation analytic study of gestures that engender repair in ESL conversational tutoring. *Journal of Pragmatics*, *42*(8), 2219–2239. https://doi.org/10.1016/j.pragma.2010.01.021

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Tominaga, W. (2013). The development of extended turns and storytelling in the Japanese oral proficiency interview. In S. J. Ross & G. Kasper (Eds.), *Assessing Second Language Pragmatics* (pp. 220–257). Palgrave Macmillan.

Youn, S. J., & Burch, A. R. (Eds.) (2020). Where conversation analysis meets langauge assessment: Toward expanding epistemologies and validity evidence (Special Issue). *Papers in Language Testing and Assessment*, *9*(1), iii–xvii. http://www.altaanz.org/uploads/5/9/0/8/5908292/2020_9_1__0_introduction.pdf

Young, R., & Milanovic, M. (1992). Discourse Variation In Oral Proficiency Interviews. *Studies in Second Language Acquisition*, *14*(4), 403–424. https://doi.org/10.1017/S0272263100011207

van Compernolle, R. A. (2013). Interactional competence and the dynamic assessment of L2 pragmatic abilities. In S. J. Ross & G. Kasper (Eds.), *Assessing Second Language Pragmatics* (pp. 327–354). Palgrave Macmillan.

Zlatev, J., Racine, T., Sinha, C., & Itkonen, I. (Eds.) (2008). *The shared mind. Perspectives on intersubjectivity*. John Benjamins.