Visual cues during interaction: Are recasts different from noncorrective repetition?

Kim McDonough, Pavel Trofimovich, Libing Lu and Dato Abashidze

McDonough, K., Trofimovich, P., Lu, L., & Abashidze, D. (2020). Visual cues during interaction: Are recasts different from noncorrective repetition? *Second Language Research*, *36*, 359–370.
https://doi.org/10.1177/0267658320914962

Abstract

Visual cues may help second language (L2) speakers perceive interactional feedback and reformulate their nontarget forms, particularly when paired with recasts, as recasts can be difficult to perceive as corrective. This study explores whether recasts have a visual signature and whether raters can perceive a recast's corrective function. Transcripts of conversations between a bilingual French–English interlocutor and L2 English university students (*N* = 24) were analyzed for recasts and noncorrective repetitions with rising and declarative intonation. Videos of those excerpts (*k* = 96) were then analyzed for the interlocutor's provision of visual cues during the recast and repetition turns, including eye gaze duration, nods, blinks, and other facial expressions (frowns, eyebrow raises). The videos were rated by 96 undergraduate university students who were randomly assigned to three viewing conditions: clear voice/clear face, clear voice/blurred face, or distorted voice/clear face. Using a 100-millimeter scale with two anchor points (0% = *he's making a comment*, 100% = *he's correcting an error*), they rated the corrective function of the interlocutors' responses while their eye gaze was tracked. Raters reliably distinguished recasts from repetitions through their ratings (although they were generally low), but not through their eye gaze behaviors.

Keywords: visual cues, eye gaze, recasts, noncorrective repetition

Visual cues during interaction: Are recasts different from noncorrective repetition?

Within the broader category of interactional feedback, recasts—which refer to an interlocutor's targetlike reformulation of a speaker's error—have attracted considerable attention due to debate about their efficacy at promoting second language (L2) learning (Goo and Mackey, 2013; Lyster and Ranta, 2013). Questions have been raised about whether L2 speakers notice recasts as being corrective, because it is their corrective function that crucially differentiates recasts from noncorrective repetition, such as a complete or partial repetition of an L2 speakers' nonproblematic utterance (Lyster, 1998; Panova and Lyster, 2002). Researchers have raised the possibility that nonverbal visual cues such as eye gaze, facial expressions, or gestures may help L2 speakers recognize interactional feedback (Gullberg, 2010; Long, 2007; Lyster, 1998). Laboratory studies have reported that recasts with gestures had a greater longer-term effect on L2 learning than recasts without gestures, although there were no advantages on the immediate posttest (Nakatsukasa, 2016). Furthermore, compared to recasts without shared gaze, recasts that occur with mutual gaze (a form of joint attention) elicit more targetlike responses in that L2 speakers produce the reformulated form (McDonough, Crowther, Kielstra and Trofimovich, 2015; McDonough, Trofimovich, Dao and Abashidze, 2018). Raters can also differentiate between questions and statements more effectively when they occur with nonverbal behaviors, although raters tend to attribute both correct and incorrect judgements to the same behaviors (Kamiya, 2018).

To further explore whether recasts occur with visual cues, researchers have asked external observers (i.e., raters) to watch and assess recast episodes. For example, Carpenter, Jeon, MacGregor and Mackey (2006) asked L2 English speakers to watch videos of recast and noncorrective repetition episodes and decide if the interlocutor was providing a correction.

Raters who heard the learners' initial utterance were more accurate at identifying recasts as corrective than raters who only heard the recast or repetition move. Stimulated recall comments from a subset of raters in both viewing conditions did not provide evidence that nonlinguistic cues played a role in the raters' decisions, as there was only one comment about facial expressions. It is possible that verbal reports are not sufficiently sensitive to identify whether raters detect and are responsive to visual cues. McDonough, Trofimovich, Lu and Abashidze (2019) recently asked raters to assess listener comprehension during nonunderstanding episodes (i.e., episodes in which the listener requested clarification) and understanding episodes (i.e., episodes in which the listener asked follow-up questions). They found that the raters could detect lack of comprehension in the nonunderstanding episodes, and their assessments were lower when raters had access to the listener's facial expressions (e.g., nods and blinks). However, these researchers focused on nonunderstanding episodes with a breakdown in the communication of meaning, so it is unclear whether visual cues would also be associated with recasts in which the interlocutor has understood the L2 speaker well enough to attempt a reformulation.

In sum, it is possible that visual cues influence how L2 speakers perceive recasts, such as by helping them recognize their corrective function. Therefore, in this study, we first examined whether recasts might be distinguished by such visual cues as head nods, frowns, blinks, eyebrow flashes, or smiles, all of which are used by speakers in communication as back-channeling signals (e.g., Bavelas, Coates and Johnson, 2002; Floyd, Manrique, Rossi and Torreira, 2016; Johnson, 1991; Knapp, Hall and Horgan, 2013). We also examined whether the corrective function of recasts is perceptible to external observers (henceforth, raters). Because we relied on an existing corpus of video-recorded interactions, it was not possible to determine how L2 speakers themselves perceive recasts. Therefore, we explored whether the corrective function of recasts would be detectable by raters from the same speech community. Although L2 speakers and raters experience the speech events under different task conditions, showing that raters can differentiate between the corrective function of recasts and noncorrective repetition raises the possibility that L2 speakers may also be sensitive to the difference. Thus, this exploratory study investigated the following two research questions:

1. Do an interlocutor's visual cues differentiate recasts from noncorrective repetitions?
2. Do raters detect differences in the corrective function of an interlocutor's recasts and noncorrective repetitions?

Method

Participants

L2 speakers. The recast episodes were drawn from a larger research project in which L2 English speakers had carried out communicative tasks with a male, bilingual French–English graduate student (McDonough et al., 2018). The 24 L2 speakers (13 women, 11 men) were enrolled in undergraduate (*n* = 12) and graduate (*n* = 11) degree programs at an English-medium university in Montreal, Canada. They ranged in age from 19 to 34, with a mean age of 25.3 years (*SD* = 4.6). They spoke four different first languages (L1s), which included Mandarin (8), Vietnamese (8), Farsi (6), and French (2), and reported having studied English for a mean of 10.4 years (*SD* = 4.3). As for their English proficiency, eight participants reported TOEFL iBT scores (*M* = 86.6, *SD* = 12.7), while the others reported IELTS scores (*M* = 6.5, *SD* = 0.4).

Raters. The raters represented the multilingual English speakers in the multicultural city of Montreal. They were recruited from the same population of university-level students as the L2 speakers, on the assumption that the raters would resemble the speakers' potential interlocutors as members of the same speech community (e.g., students enrolled in the same courses). The 96 undergraduate students (60 women, 36 men), who were studying in a variety of degree programs at the same university, ranged in age from 18 to 57 (M = 23.5, SD = 6.6). Almost all raters were multilinguals (n = 71) or bilinguals (n = 24), with only one English monolingual. Fifty-eight (59%) of the raters reported either English or French as an L1, with the other raters speaking 17 other L1s. Those with a non-English L1 self-evaluated their English proficiency using a 9-point Likert scale (1 = *beginner*, 9 = *nativelike*), with a mean ranking of 7.5 (SD = 1.1). In terms of these raters' familiarity with accented English, their mean score on a 9-point Likert scale (1 = *not at all familiar*, 9 = *very familiar*) was 7.1 (SD = 1.9). L1 English raters self-assessed their English proficiency at a mean ranking of 8.9 (SD = 0.4) and estimated their familiarity with accented English at a mean of 7.3 (SD = 1.9). All raters had previously taken language courses, ranging from English writing or foreign languages classes to linguistic courses in phonology or grammar.

Materials and Procedure

Episode selection. The interaction data had been collected for a larger study in which L2 speakers carried out communicative tasks with a research assistant (i.e., the interlocutor) while their eye gaze was video-recorded using the faceLab 5.0 system (McDonough et al., 2018). As part of the larger study, audio recordings of their interaction had been transcribed and verified by research assistants. For the current study, the third author analyzed the transcripts to identify recasts and noncorrective repetitions of similar length from the same L2 speaker. All episodes were chosen based on the transcripts exclusively, without any prior knowledge of whether the videos would contain visual cues. Whereas a recast was the interlocutor's reformulation of an L2 speaker's error, noncorrective repetition was the interlocutor's verbatim repetition (partial or complete) of an L2 speaker's utterance without any corrections. Thus, the crucial distinction was whether an error had occurred in the L2 speaker's initial utterance, as both response types can serve the same discourse function of moving a conversation forward. We included an equal number of recasts and noncorrective repetitions with interrogative and declarative intonation from each speaker. Each L2 speaker contributed two recasts and two noncorrective repetitions to the dataset, which yielded a total of 48 recast and 48 noncorrective repetition episodes. Sample recasts and noncorrective repetitions (both in Turn 2 of the interlocutor's speech) appear in Table 1.

Table 1. *Examples of recasts and noncorrective repetitions*

Turn	Recast	Noncorrective repetition
1	S: Have you have any dream of going to moon?	S: Ah ok so you are not originally from Quebec yeah?
2	I: To the moon?	I: From Quebec?
3	S: Yeah something like going to the space and such a technology	S: Yeah
4	I: Well I think I would just like to take a picture of Earth but actually going, not sure.	I: Well I mean yes I am. I am from Quebec. It's just that my family is bilingual.

Note: S = L2 Speaker; I = Interlocutor

Rating stimuli. After selecting the 96 episodes, we edited the videos to show the interlocutor's image (from upper shoulders to face) while he was providing a recast or noncorrective repetition (i.e., Turn 2 in sample sequences illustrated in Table 1). The videos did not show his arms or hands so that any gestures were not visible, leaving his face as the focus of the viewing frame. The L2 speakers' initial utterances were not included in the videos because access to that utterance would reveal to the raters whether an error had been produced, thereby helping them assess corrective function. To ensure variation in the raters' access to auditory and visual information, three versions of each video were created: clear face and clear voice (+*face*/+*voice*), blurred face with clear voice (–*face*/+*voice*), and clear face with distorted voice (+*face*/–*voice*). The different versions could help determine whether access to visual information affects raters' ability to assess an utterance's corrective function. The –*face* versions were created by blurring the interlocutor's face using the "mosaic" function in Adobe Premiere Pro. A "square block" set at 10° of visual angle was placed from the top of the interlocutor's forehead to his chin. The –*voice* versions were created through low- and high-pass filtering within the range of 400–23,000 Hz and were edited through the "pitch shifter" function by setting up a semi-tone (temporal setting of –8 milliseconds, at 350 Hz frequency, and a 50% overlap). Videos in each condition, which was a between-groups variable, were randomized into four lists. All videos were presented using the EyeWorks 3.0 program where raters clicked on the "continue" button to go to the next video whenever they were ready.

Rating procedure. The 96 raters were randomly assigned to one of the three video conditions: +face/+voice (32), –face/+voice (31), and +face/–voice (33). They watched the videos and evaluated the corrective function of the interlocutor's utterances during an individual 75-minute session. After filling out consent and background information forms (15 minutes), they read the instructions and definitions of key terms in the rating booklet, reviewed the rating scale, and asked any clarification questions (10 minutes). The rating booklet defined corrective function as the rater's perception about whether the interlocutor's utterance had corrected an error. The rating scale was a continuous 100-millimeter scale with two descriptive anchor points: 0% = *he's making a comment* and 100% = *he's correcting an error*. After reviewing the definitions and rating scale, the raters then used the scale to rate two practice videos and 96 target videos (30 minutes). After completing the videos, the raters participated in a debriefing session (20 minutes) in which they answered questions about their experience rating the videos and provided additional ratings with commentary for four previously unseen videos (two recasts and two noncorrective repetitions) to elicit their perceptions about the rating process.

Data Analysis

The videos were analyzed for the occurrence of visual cues following a bottom-up, data-driven approach because we had no a priori expectations as to which cues might be associated with recasts versus noncorrective repetitions. The third author watched all the videos to identify facial expressions, which were classified into the following categories: (a) head nods, (b) blinks, (c) upper facial expressions (looks away from the speaker, raised eyebrows), and (d) lower facial expressions (smiling, frowning, pursed lips, open mouth). After establishing the categories, she watched the videos again to record frequency counts for each category, after which a peer reviewed the videos and frequency counts with any disagreements resolved through discussion. A research assistant independently then coded all the videos, and interrater reliability, assessed using two-way random average-measures intraclass correlation coefficients, was .97 for nods, .92 for blinks, and .85 for facial expressions. Mean occurrence of facial expressions was obtained by dividing the sum in each category by the total numbers of recast and repetition episodes separately.

For visual cues in the form of eye gaze, the interlocutor's eye gaze during the recasts and noncorrective repetitions had been recorded in milliseconds using the EyeWorks 3.0 program (http://www.eyetracking.com/Software/EyeWorks) as part of the larger study. Using the analysis program Captiv (http://teaergo.com/wp/tea-behavior-analyses-products/captiv-solution-2/?lang=en), the interlocutor's fixations during the recast and repetition turns were manually coded frame-by-frame to determine in which area of the scene he had fixated, and the total eye gaze to the L2 speaker's face was summed. We acknowledge that this analysis necessarily targeted the interlocutor's overt attention to visual cues through eye fixations and that the interlocutor may have captured additional visual cues through peripheral vision (e.g., Gullberg and Holmqvist, 2006). Nevertheless, for this exploratory study, a focus on the interlocutor's overt attention to visual cues was deemed sufficient to provide initial insights into a speaker's visual behavior during recasts and noncorrective repetitions.

To account for variation in turn length, interlocutor eye gaze duration was calculated as a proportion of total turn length (eye gaze duration to L2 speaker's face/turn duration). Duration was calculated beginning the moment that the interlocutor's gaze first landed on the face of the L2 speaker and ended the moment his gaze left the speaker's face. For example, Figure 1 shows the interlocutor's eye gaze to the L2 speaker during a response turn. The green dot represents where he was looking and the line shows the movement of the dot across his field of vision. The eye gaze duration was the time that the green dot was located on the L2 speaker's face divided by the total speaking time for his utterance.

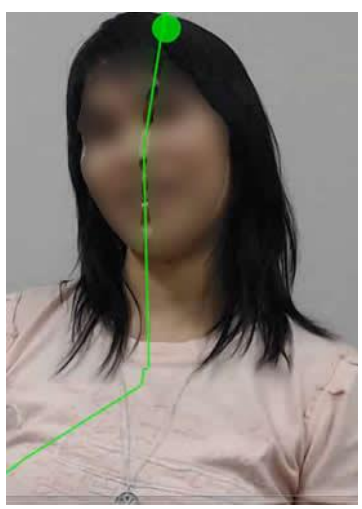

Figure 1 *Interlocutor eye gaze during response turn*

A subset of the eye gaze data (20%) was coded by a research assistant to obtain interrater reliability using a two-way random average-measures intraclass correlation coefficient, which was .99. For the raters' assessments of corrective function, the location of their checkmarks or crosses on the scale was converted to numeric score (0–100) by measuring the distance from the leftmost edge of the 100-millimeter scale. Alpha was set at .05 for all statistical tests.

Results

The first research question asked whether an interlocutor's visual cues differentiate recasts from noncorrective repetitions. As shown in Table 2, the interlocutor's eye gaze durations during recasts and noncorrective repetitions were similar as he spent approximately one-third of his turn (measured in seconds) looking at the L2 speakers' faces. For the other visual cues, head nods, blinks, and facial expressions were all more frequent during noncorrective repetitions. In light of the small sample size and lack of normal distribution, the values were compared using separate Wilcoxon signed-ranks tests, which are the nonparametric equivalent of paired-samples *t* tests. Only differences in upper facial expressions reached statistical significance, which included raised eyebrows and looks away from the speaker, such as directing eye gaze downward toward the table or to the left and right edges of the field of vision. Thus, the results for the first question indicate that facial expressions in the interlocutor's upper face (eye) region differentiated between the recasts and noncorrective repetitions, occurring more frequently with noncorrective repetitions.

Table 2. *Interlocutor's mean visual cues per recast/noncorrective repetition turn*

	Recast		Repetition		Comparison		
Visual cue	*M*	*SD*	*M*	*SD*	*Z*	*p*	*d*
Eye gaze duration (proportion of total turn time)	0.33	0.39	0.36	0.33	0.33	.744	0.08
Head nods	0.96	1.23	1.38	1.10	1.16	.224	0.36
Blinks	0.63	0.92	0.83	0.76	1.07	.284	0.24
Upper facial expressions	0.54	0.59	1.08	0.97	2.04	.042	0.67
Lower facial expressions	0.38	0.58	0.63	0.65	1.60	.109	0.41

The second research question asked whether external observers (i.e., raters) detect differences in the corrective function of the interlocutor's recasts and noncorrective repetitions. The mean ratings from the three video rating conditions appear in Table 3 by episode type. The descriptive trends were that recasts received higher corrective function ratings in all three rating conditions, with a larger difference in mean scores when raters had access to both the interlocutor's voice and face. A mixed ANOVA with one between-groups variable (rating condition) and one within-groups variable (episode type) indicated that there was a significant main effect for episode type, $F(1, 93) = 29.24$, $p = .001$, partial $\eta^2 = .24$, but no significant main effect for rating condition, $F(2, 93) = 0.66$, $p = .518$, partial $\eta^2 = .01$, and no significant two-way interaction, $F(2, 93) = 1.55$, $p = .218$, partial $\eta^2 = .03$. Raters attributed significantly higher corrective function ($d = 2.29$) to recasts ($M = 40.31$, $SD = 1.58$) than noncorrective repetitions ($M = 36.75$, $SD = 1.53$), seemingly through the use of either visual (face) or auditory (voice) information available from the interlocutor (i.e., without the additive effect of having access to the interlocutor's face).

Table 3. *Corrective intent ratings (out of 100) by video condition and episode type*

	Recast			Repetition		
Rating condition	*Min–Max*	*M*	*SD*	*Min–Max*	*M*	*SD*
+face/+voice (n = 32)	8.80–83.00	43.54	16.89	2.20–65.20	38.34	14.86
–face/+voice (n = 31)	9.80–65.10	39.10	15.39	6.50–68.70	36.47	17.15
+face/–voice (n = 33)	8.80–68.90	38.30	14.07	8.70–62.90	35.43	12.88

Discussion

This exploratory study examined whether recasts and noncorrective repetitions are associated with distinct visual cues provided to L2 speakers by their interlocutors, and whether observers of interactions distinguish between the corrective function of recasts and noncorrective repetitions. Results revealed that the interlocutor tended to show more expressions in the upper part of the face (downward looks, raised eyebrows) when providing L2 speakers with noncorrective repetitions than when recasting the speakers' erroneous utterances, which may reflect his reactions to the content of the L2 speaker's utterance when asking follow-up questions. When raters observed the interlocutor, they reliably distinguished recasts from repetitions through their ratings. Although their ratings were on the lower end of the 100-point scale, raters perceived recasts as having a greater corrective function than repetitions.

Our first objective was to determine if recasting and noncorrective repetition would be associated with distinct visual signatures for the interlocutor, on the assumption that the two conversational moves might be distinguished through visual cues provided to L2 speakers by their interlocutor. Previous authors have reported clear differences between visual cues signaling interlocutors' understanding of a speaker's utterance and those indicating a communication breakdown (McDonough et al., 2019; Floyd et al., 2016). Unlike understanding episodes, communication breakdowns are typically characterized by holds, which refer to temporary cessation of all dynamic movements involving gaze, head position, upper body lean, or eyebrows, with release occurring during or shortly after the speaker provides clarification, with holds possibly triggered by an increase in the processing load created by nonunderstanding. Because neither recasts nor noncorrective repetitions involve a communication breakdown or lack of intelligibility, it may not be altogether surprising that these conversational moves did not elicit robust differences in visual cues in this dataset. Nevertheless, that noncorrective repetition co-occurred more frequently with upper facial expressions such as downward looks and raised eyebrows is an intriguing finding which must be revisited systematically in future research to determine whether it is a general pattern across interlocutors. It will also be important to explore whether more explicit forms of feedback also evoke holds, due to a greater interruption to communicative flow than recasts, or whether holds occur exclusively when there is a breakdown in the communication of meaning.

One reason for why recasting and noncorrective repetition may not have elicited compelling visual signatures is that visual cues might precede the recasts rather than co-occur with them, such that any visual cues associated with recasting were provided by the interlocutor before the recast when he first detected the L2 speaker's problematic form. To determine whether visual cues precede recasts, thereby potentially signalling what follows as corrective, we carried out a post hoc analysis of the interlocutor's visual cues and eye gaze duration during the L2 speakers' initial turn. Although holds were more frequent in the recast episodes (5/24) than in the repetition episodes (1/24), a chi-square test with a continuity correction indicated that the association between episode type and holds was not significant, $\chi^2(1, 48) = 1.71$, $p = .190$, Cramer's $V = .25$. As shown in Table 4, the interlocutor provided head nods and blinks more frequently before recasts than before noncorrective repetitions. However, the interlocutor looked at the speaker longer and engaged in more upper and lower facial expressions before repetitions. Only the difference for blinks reached statistical significance, with a Cohen's *d* effect size of 0.49 (see Table 4). Incidentally, in our own prior work, blinks were more frequent in nonunderstanding episodes than understanding episodes (McDonough et al., 2019). Given that blinks are used to initiate a repair (e.g., Johnson, 1991) and are linked to affective arousal and cognitive difficulty (Knapp, Cody and Reardon, 1987; see Eckstein, Guerra-Carrillo, Singley and Bunge, 2017, for a discussion on blink rate and processing), an interlocutor's blinking that co-occurs with a speaker's nontarget utterance might be an important cue signaling that repair is forthcoming and/or that the interlocutor is having difficulty understanding.

Table 4. *Interlocutor's mean visual cues per turn preceding recasts/noncorrective repetitions*

	Recast		Repetition		Comparison		
Visual cue	*M*	*SD*	*M*	*SD*	*Z*	*p*	*d*
Eye gaze duration (proportion of total turn time)	0.44	0.37	0.50	0.40	1.49	.136	0.16
Head nods	1.50	1.75	0.88	1.51	1.37	.172	0.38
Blinks	3.83	1.74	2.88	1.15	2.15	.031	0.49
Upper facial expressions	0.63	0.88	0.75	0.94	0.83	.405	0.13
Lower facial expressions	1.71	0.75	1.83	1.17	0.36	.718	0.12

Our second objective was to examine whether external observers can distinguish the corrective function of an interlocutor's recasts and noncorrective repetitions. By providing different rater groups with videos containing both the interlocutor's image and his utterance (+face/+voice), only the interlocutor's image (+face/–voice), or only his utterance (–face/+voice), we sought to determine if access to the interlocutor's face influences raters' assessment of corrective function. However, because raters attributed higher scores to recasts than to noncorrective repetitions irrespective of the rating condition, it appears that corrective intent can be cued through multiple information channels, either visual or auditory, with visual information alone providing no additional benefit for raters to distinguish between corrective and noncorrective repetitions, which contrasts with generally positive effects of multimodal (audiovisual) training for specific L2 skills (e.g., Godfroid, Lin and Ryu, 2017). However, it is not known whether L2 speakers themselves rely on both auditory and visual information to interpret response turns during real-time conversation, which is a question we are currently investigating.

While it is unclear which specific auditory and visual details raters used in their evaluation of the interlocutor's corrective intent, the ratings were low overall (see Table 3), falling in the 35–45 mean range on a 100-point scale across the three conditions. During the debrief interview, most raters also commented on their difficulty with rating the interlocutor's corrective intent, citing the need for more information (e.g., more context or longer video episodes) to provide confident judgments. Raters also mentioned the same interlocutor behaviors (e.g., looking confused) as cues for both recasting and noncorrective repetition or cited one behavior (e.g., nodding) as being uniquely associated with one conversational move (i.e., noncorrective repetition) although this behavior occurred in both episode types. Taken together, these findings imply that it was generally difficult for raters to identify corrective function, which is in line with prior work showing that L2 speakers often fail to notice recasts (e.g., Carpenter et al., 2006; Lyster, 1998; Lyster and Ranta, 1997; Panova and Lyster, 2002).

The current findings must remain speculative until they are revisited in future research and replicated with a variety of interlocutors, preferably using sensitive eye-tracking measures of both focal and peripheral visual attention. Here, the links between the interlocutor's visual cues and different response types were tenuous. Any visual signature of recasting that makes it distinct from noncorrective repetition might manifest itself more strongly in classroom-based interaction, perhaps particularly when recasting is paired with other nonverbal cues such as gestures (Nakatsukasa, 2016) or when an instructor purposefully uses visual cues to enhance corrective feedback when focusing on form (Davies, 2006; for an example, see Meek, 2010). In future classroom-based studies, it would be useful to explore whether students are sensitive to their instructor's use of nonverbal cues when interacting with their classmates, such as when an instructor indicates nonunderstanding or provides feedback. Just as "observers" of interaction might possibly benefit without directly participating in the conversation (cf. Mackey, 1999), students who observe instructor–student interactions may also rely on nonverbal cues when interpreting instructor responses.

Crucially, future research must clarify the extent to which any visual cues made available by interlocutors concurrently with repetitions (corrective or otherwise) are interlocutor-specific and to which they are subject to cultural differences, as both visual cues and eye gaze behaviors vary inter-personally and cross-culturally (e.g., Rossano et al., 2009; Vatikiotis-Bateson et al., 1998; Zhang and Kalinowski, 2012). In light of their greater power (Godfroid, 2020), within-groups designs should further probe the individual and combined impact of auditory and visual information on external observers' ability to differentiate among an interlocutor's response types.

Similarly, it would be important to focus on L2 speakers themselves, as opposed to external observers, investigating the extent to which they are able to detect and act on an interlocutor's visual cues. Researchers might also consider moving away from recasting by targeting other interactional moves such as confirmation checks or clarification requests, with the view of understanding whether such moves are associated with visual cues that could be harnessed pedagogically so as to enable L2 speakers to avoid communication difficulties and initiate self-repair. Finally, future research targeting L2 speakers' use of visual cues during interaction must be conceptualized theoretically, ideally within frameworks encompassing both linguistic and pragmatic development (for recent examples, see Culpeper, Mackey and Taguchi, 2018, and Taguchi and Roever, 2017), so that the role of visual cues in L2 use can be problematized and understood not just descriptively but as part of broader views of language use and language development.

References

Bavelas J, Coates L and Johnson T (2002) Listener responses as collaborative process: The role of gaze. *Journal of Communication* 52: 566–580. https://doi.org/10.1111/j.1460-2466.2002.tb02562.x

Carpenter H, Jeon KS, MacGregor D and Mackey A. (2006) Learners' interpretation of recasts. *Studies in Second Language Acquisition* 28: 209–236. https://doi.org/10.1017/S0272263106060104

Culpeper J, Mackey A and Taguchi N (2018) *Second Language Pragmatics: From Theory to Research*. New York: Routledge. https://doi.org/10.4324/9781315692388

Davies M (2006) Paralinguistic focus on form. *TESOL Quarterly* 40: 841–855. https://doi.org/10.2307/40264316

Eckstein MK, Guerra-Carrillo B, Singley, ATM and Bunge, SA (2017) Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience* 25: 69–91. https://doi.org/10.1016/j.dcn.2016.11.001

Floyd S, Manrique E, Rossi G and Francisco T (2016) Timing of visual bodily behavior in repair sequences: Evidence from three languages. *Discourse Processes* 53: 175–204. https://doi.org/10.1080/0163853X.2014.992680

Godfroid A (2020) *Eye-Tracking in Second Language Acquisition and Bilingualism: A Research Synthesis and Methodological Guide*. New York: Routledge.

Godfroid A, Lin C and Ryu C (2017) Hearing and seeing tone through color: An efficacy study of web-based, multimodal Chinese tone perception training. *Language Learning* 67: 819–857. https://doi.org/10.1111/lang.12246

Goo J and Mackey A (2013) The case against the case against recasts. *Studies in Second Language Acquisition* 35: 127–165. https://doi.org/10.1017/S0272263112000708

Gullberg M (2010) Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research* 26: 75–102. https://doi.org/10.1177/0267658309337639

Gullberg M and Holmqvist K (2006) What speakers do and what addressees look at: Visual attention to gestures in human interaction live and on video. *Pragmatics & Cognition* 14: 53–82. https://doi.org/10.1075/pc.14.1.05gul

Johnson K (1991) Miscommunication in interpreted classroom interaction. *Sign Language Studies* 70: 1–34. https://doi.org/10.1353/sls.1991.0005

Kamiya N (2018) The effect of learner age on the interpretation of the nonverbal behaviors of teachers and other students in identifying questions in the L2 classroom. *Language Teaching Research* 22: 47–64. https://doi.org/10.1177/1362168816658303

Knapp ML, Cody MJ and Reardon KK (1987) Nonverbal signals. In: Berger C and Chafee S (eds) *Handbook of Communication Science*. Newbury Park: Sage, pp. 385–418.

Knapp ML, Hall JA and Horgan TG (2013) *Nonverbal Communication in Human Interaction*. Boston: Wadsworth.

Long M (2007) *Problems in SLA*. Mahwah: Lawrence Erlbaum.

Lyster R (1998) Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition* 20: 51–81. https://doi.org/10.1017/S027226319800103X

Lyster R and Ranta L (2013) Counterpoint piece: The case for variety in corrective feedback research. *Studies in Second Language Acquisition* 35: 167–184. https://doi.org/10.1017/S027226311200071X

Mackey A (1999) Input, interaction, and second language development: An empirical study of question formation in ESL *Studies in Second Language Acquisition* 21: 557–587. https://doi.org/10.1017/S0272263199004027

McDonough K, Crowther D, Kielstra P and Trofimovich P (2015) Exploring the potential role of eye gaze in eliciting English L2 speakers' responses to recasts. *Second Language Research* 31: 563–575. https://doi.org/10.1177/0267658315589656

McDonough K, Trofimovich P, Dao P and Abashidze D (2018) Eye gaze and L2 speakers' responses to recasts: A systematic replication study of McDonough, Crowther, Kielstra, and Trofimovich (2015). *Language Teaching.* Advance online publication. https://doi.org/10.1017/S0261444818000368

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

Meek M (2010, March 1) Correcting behavior with hand signals. *The New York Times Magazine*. Retrieved from https://www.nytimes.com/video/multimedia/1247467223003/correcting-behavior-with-hand-signals.html

Nakatsukasa K (2016) Efficacy of recasts and gestures on the acquisition of locative prepositions. *Studies in Second Language Acquisition* 38: 771–799. https://doi.org/10.1017/S0272263115000467

Panova I and Lyster R (2002) Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly* 36: 573–595. https://doi.org/10.2307/3588241

Rossano F, Brown P and Levinson SC (2009) Gaze, questioning and culture. In: Sidnell J (ed) *Conversation Analysis: Comparative Perspectives*. Cambridge: Cambridge University Press, pp. 187–249. https://doi.org/10.1017/CBO9780511635670.008

Taguchi N and Roever C (2017) *Second Language Pragmatics*. Oxford: Oxford University Press.

Vatikiotis-Bateson E, Eigsti IM, Yano S and Munhall KS (1998) Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics* 60: 926–940. https://doi.org/10.3758/BF03211929

Zhang J and Kalinowski J (2012) Culture and listeners' gaze responses to stuttering. *International Journal of Language and Communication Disorders* 47: 388–397. https://doi.org/10.1111/j.1460-6984.2012.00152.x