Studies in Second Language Acquisition, 2014, **36**, 1–37.
doi:10.1017/S0272263113000363

EXPLORING LEARNER PERCEPTION AND USE OF TASK-BASED INTERACTIONAL FEEDBACK IN FTF AND CMC MODES

Laura Gurzynski-Weiss
Indiana University

Melissa Baralt
Florida International University

Theoretical claims about the benefits of corrective feedback have been largely premised on learners' noticing of feedback (e.g., Gass & Mackey, 2006; Long, 1996; Schmidt, 1990, 1995; Swain, 1995), and findings have demonstrated that both the feedback target (Mackey, Gass, & McDonough, 2000) and the mode of provision (Lai & Zhao, 2006) can affect learners' accurate perception of feedback. The current study extended this research by investigating learners' perception and use of feedback provided in task-based interaction in both computer-mediated (CMC) and face-to-face (FTF) modes. Utilizing stimulated recall, the study examined if 24 intermediate-level learners of Spanish as a foreign language accurately noticed feedback as feedback, if they noticed the feedback target, and if the environment in which they interacted (CMC vs. FTF) made a difference in their accuracy. The study also investigated if modality affected opportunities for modified output immediately following feedback and if learners used those opportunities differently according to mode. Results demonstrated that, overall, learners did notice feedback as feedback in both modes. Contrary to expectations, there were no statistical differences between modes

The authors would like to express their appreciation to the students who participated in this study. An earlier version of this article was presented at the Georgetown University Roundtable on Linguistics in March, 2009.

Correspondence concerning this article should be addressed to Laura Gurzynski-Weiss, 1020 E. Kirkwood Ave., Ballantine 876, Bloomington, IN 47405. E-mail: lgurzyns@indiana.edu

© Cambridge University Press 2013

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

in feedback perception accuracy. Significant differences were found, however, in learners' opportunities for and use of feedback depending on the interaction environment and the type of error being addressed.

For corrective feedback provided during task-based interaction to be effective, learners have to notice it (e.g., Schmidt, 1990, 1995). This critical noticing is dependent not only on learners' awareness that the feedback was corrective in nature but also on their noticing the target of that feedback. Mackey, Gass, and McDonough (2000) were the first to use stimulated recall methodology to explore whether learners accurately perceive interactional feedback. They found that learners were generally accurate about feedback that addressed lexical, semantic, and phonological errors but not about feedback that addressed morphosyntactic errors. Lai and Zhao (2006) extended Mackey et al.'s methodology to the computer-mediated communication (CMC) environment and compared learners' perception of feedback in a face-to-face (FTF) environment to the CMC mode. They reported that CMC may enhance learners' noticing of feedback, but that this could depend on the type of feedback employed. The two studies raised key questions about which feedback types are most effective for specific errors and whether noticing of feedback may be differentially experienced depending on the environment in which learners interact. These studies, however, examined only two types of feedback: recasts and negotiation of meaning (and a combination of the two, in the case of Mackey et al.). The present study aims to test the external validity of these two experiments by using stimulated recall protocols to explore learners' noticing of various forms of corrective feedback in a different context: that of Spanish foreign language learning in both CMC and FTF task-based interaction. Additionally, the current study extends this line of research by examining the opportunity to use the feedback, as well as learners' incorporation of feedback (or lack thereof), to assist in addressing the hypothesis that learners must notice feedback in order to benefit from it.

BACKGROUND

The Role of Second Language Corrective Feedback

According to Schmidt's (1990, 1995) noticing hypothesis, learners must consciously notice linguistic forms and the meanings that they carry for second language (L2) acquisition to take place. This role for noticing is the very foundation on which the efficacy of corrective feedback, or reactions to learners' erroneous production, is based. The interaction

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

hypothesis (e.g., Gass, 1997; Gass & Mackey, 2006; Long, 1996, 2007) ascribes a key role for corrective feedback in interaction because of its potential to help the learner notice the mismatch between his or her nontargetlike production and the target form. Highlighting the benefits of conversational interaction, Long (1996) argued that the negotiation work that occurs during interaction between learners and interlocutors is what productively connects input, noticing of that input, and output. Although learners' modified output as a response to feedback cannot be equated with development (e.g., Egi, 2010; Mackey & Philp, 1998), it may shed light on how learners perceive the feedback. For example, Swain (1995) stated that learners' modified output can be indicative of "the leading edge of a learner's interlanguage" (p. 131).

With five recent meta-analyses demonstrating empirically that feedback is beneficial for learning (Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006; Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007; Russell & Spada, 2006), the issue of whether or not corrective feedback plays a causal role in L2 acquisition is no longer a contentious one in the field. Researchers are now asking the more fundamental question of which factors moderate the effectiveness of feedback; these include (a) cognitive and affective learner-internal variables, such as why feedback is more effective for some learners and less so for others (e.g., Goo, 2012; Iwashita, 2003; Mackey & Philp, 1998; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Philp, 2003; Sheen, 2007; Tarone & Bigelow 2007; Trofimovich, Ammar, & Gatbonton, 2007); (b) external factors, such as the design of the task used during task-based interaction (e.g., Baralt, in press; Gurzynski-Weiss & Révész, 2012; Révész, 2009; Révész & Han, 2006); and (c) characteristics of the providers of feedback (i.e., the instructor or interlocutor) and the manner in which their background variables mediate their decision to provide feedback or not (e.g., Gurzynski-Weiss, 2010, in press). A rigorous line of research on L2 feedback has looked at (a) the nature of feedback itself, such as prosodic features and overall length (e.g., Loewen & Philp, 2006); (b) how many changes a feedback episode makes to the erroneous utterance (e.g., Philp, 2003); (c) the linguistic target (e.g., Havranek, 2002; Jeon, 2007); (d) whether the feedback provides, repeats, or elicits the correct form to or from learners (e.g., Ammar & Spada, 2006; Carpenter, Jeon, MacGregor, & Mackey, 2006; Lyster, 2004; Nassaji, 2007); (e) what features render feedback more implicit versus explicit (see Ellis, 2009, for a review); and, most important for the current study, (f) how learners respond to different types of feedback (e.g., Lyster, 2001; Lyster & Ranta, 1997). All of this empirical work is, of course, centered on the assumption that learners notice feedback as feedback.

Ascertaining learners' noticing of feedback, and, if they do notice it, whether they accurately notice the target of the feedback, is challenging from a methodological standpoint. This is especially the case for exploring noticing of corrective feedback that is provided during

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

two-way conversational interaction. Examining what learners do with feedback is one means by which researchers have attempted to shed light on its effectiveness (e.g., Egi, 2010). For example, if a learner incorporates the feedback, thus modifying his or her original, nontargetlike production, it could mean that he or she has noticed the feedback. From a theoretical standpoint, Swain (1995) has highlighted that modified output immediately after feedback facilitates the language learning process because it is synonymous with a reprocessing of the erroneous output, which leads to further hypothesis testing about the correct form and can result in maintenance of those modifications in their interlanguage. A handful of studies have shown a relationship between learners' production of modified output after feedback and subsequent learning (e.g., Havranek, 2002; Loewen, 2005). At the same time, many have not, reporting instead that learners' lack of a response could be a "red herring" (Mackey & Philp, 1998, p. 338) or that their immediate use of feedback was nothing more than a parroting of that feedback without conscious noticing that the feedback was corrective in nature (e.g., Egi, 2010; Mackey & Philp, 1998; McDonough & Mackey, 2006; Révész, Sachs, & Mackey, 2011).

As explained by Long (2007), another problem with examining learner responses to feedback as being indicative of noticing is that "discourse constraints often preempt the possibility of repair in the immediately following turn, especially in classroom settings" (p. 94). That is, the opportunity to use feedback in interactive discourse is not always present. Mackey and Philp (1998) addressed this question empirically. They reported that 22% of recasts were categorized as "no opportunity turns" (p. 349), or recasts that were impossible to repeat or incorporate. Therefore, learners' incorporation of feedback may be dependent not only on their noticing of that feedback but also on whether or not they are given an opportunity to use it. So far, minimal research has explored opportunities for modified output in relation to produced modified output and measures of noticing.

Stimulated Recall to Measure Noticing of Feedback

One effective means to infer learner noticing of feedback during task-based interaction is through stimulated recall (see Gass & Mackey, 2000, for a detailed review). Stimulated recall (SR) is a retrospective procedure during which the learner is shown a stimulus (e.g., a video of him- or herself engaging in interaction) to prompt the learner's accurate memory of what occurred at the time of the original interaction. During the procedure, the researcher stops the video after feedback episodes and elicits a verbalization of the learner's thought processes at the time of the episode. This is done by asking the learner what he or she was thinking at that moment; the learner may also be invited to stop the

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

video (or other stimulus) to reflect and add an additional comment. To reduce the potential of memory decay and improve the robustness of the procedure, the retrospection should be carried out immediately after the task. The frequent use of this procedure in interaction research attests to its validity (e.g., Carpenter et al., 2006; Egi, 2010; Gass & Lewis, 2007; Swain & Lapkin, 2002).

Mackey et al. (2000) were the first in the field of SLA to employ SR methodology to examine learner perception of corrective feedback during task-based interaction. Ten beginning- and intermediate-level learners of English as a second language (ESL) as well as seven Italian as a foreign language (IFL) learners participated in the study. The participants carried out a spot-the-difference picture task with a native ESL or near-native IFL interlocutor. During the interaction, the interlocutor responded to nontargetlike utterances with three types of feedback: (a) recasts, (b) negotiation of meaning, or (c) a combination of the two, whenever the interlocutor felt it was natural and appropriate to do so (p. 479). These feedback types were provided in response to learners' phonological, morphosyntactic, lexical, and semantic errors. After the task, the researchers showed each learner the video of their interaction and inquired about what their thoughts were after having received feedback. On the basis of what the participants reported, the researchers then categorized participants' SR comments as *lexical*, *semantic*, *phonological*, *morphosyntactic*, *no content*, or *unclassifiable*. They found that most of the feedback received by ESL learners addressed phonological and then morphosyntactic errors. However, learners often incorrectly perceived feedback as targeting a different linguistic structure than intended by their interlocutor. For example, ESL learners only recognized feedback targeting morphosyntactic errors as such 13% of the time; rather, they identified it as targeting semantics (38% of the time) or about nothing in particular (21% of the time). These learners were most accurate in their perception of lexical (83%) and phonological feedback (60%). Looking at IFL learners, the researchers found that the feedback they received addressed mostly lexical and then morphosyntactic errors; there were fewer phonological feedback episodes. These learners were most accurate about lexical feedback (66%), whereas they perceived morphosyntactic feedback accurately only 24% of the time. In sum, both groups of participants were generally accurate about their perceptions of lexical and phonological feedback but not about feedback addressing morphosyntax. On the basis of this finding, the researchers raised the important question of the utility of interactional feedback addressing morphosyntax. In a post hoc analysis that examined the type of feedback given for error types and learner-produced uptake (defined as the learners' modification of their original utterance immediately following feedback), Mackey et al. (2000) found that recasts were the most common type of feedback that addressed morphosyntactic errors (i.e., 75% of the time).

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Negotiation of meaning or a combination of recasts and negotiation of meaning were the most frequent feedback types addressing other error types. The researchers asked the critical question of whether recasts after morphosyntactic errors are "suboptimal," given that learners in their study did not perceive them accurately most of the time (p. 493). They suggest that this may be due to the fact that accurate morphosyntactic production was not critical for comprehension during the specific task used in their study.

Interaction in CMC: A Way to Promote Noticing?

Due to the increase in the development of online and distance language classes (see Blake, 2011, for a review), SLA researchers and practitioners have begun to explore how the online as compared to the FTF environment may differentially mediate noticing and subsequent language acquisition. Some have posited that learners may notice feedback more in CMC, given that its written mode may increase the saliency of forms addressed in corrective feedback (e.g., Chapelle, 2001; Long, 2007). Sauro (2009) has also highlighted the cognitive relief that CMC may provide because, in CMC, learners do not have to compare feedback in memory; rather, their original nontargetlike utterance and the corrected utterance are stored and accessible throughout the exchange. Interaction in CMC also involves slower turn taking, which obviates the social need to have to respond immediately and may give learners extra online planning time to construct their output. These characteristics of the CMC environment are hypothesized to promote more noticing of form during interaction (see Ortega, 2009, for a review).

Lai and Zhao (2006) were the first researchers to use SR to examine learners' noticing and perception of feedback in the CMC mode, comparing it to noticing in FTF mode. Unlike Mackey et al. (2000), Lai and Zhao did not use native speakers or expert interlocutors as dyadic partners. Instead, 12 learners of ESL were paired into mixed-proficiency dyads with each other (six high-intermediate- and six low-intermediate-level learners, as determined by a placement test). Participants carried out two spot-the-difference picture tasks, one in FTF mode and one in CMC. The stimulus used for the CMC interaction was a screen-recording video of the chat interaction. As in Mackey et al., the researchers aimed to explore whether learners noticed feedback (negotiation of meaning or recasts) accurately as feedback, which they measured by comments during SR protocols. Lai and Zhao also examined whether participants noticed their own errors, which they operationalized as learners' self-corrections during the interaction.

Lai and Zhao (2006) found that learners noticed negotiation for meaning significantly more often in CMC-based interaction, as compared to FTF

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

(45% compared to 24%, respectively), even though significantly more negotiation of meaning occurred in FTF interaction. There was, however, no statistical difference in learners' noticing of recasts (10% in CMC compared to 18% in FTF). Examining the distribution of recasts addressing different error types in each mode, Lai and Zhao reported that 65% of all recasts provided in CMC addressed morphosyntactic errors, whereas only 28% of recasts addressed morphosyntax in FTF mode. Instead, and counter to Mackey et al.'s findings, recasts mostly addressed lexis in the FTF mode (39%, compared to CMC's 17%). The researchers also reported on learners' opportunities for modified output after receiving recasts. Both modes resulted in nearly identical opportunities: Learners had opportunities to modify their output after 70% of recasts in CMC and after 71% in FTF. In examining learners' noticing of their own errors, the CMC mode resulted in significantly more self-corrections than FTF. Lai and Zhao cited the additional processing time that learners have as well as the "permanency of the text" as reasons why CMC enhanced noticing of feedback (p. 112). Thus, according to their findings, learners may notice corrective feedback (in the form of negotiation of meaning) as well as their own errors better in CMC mode than they do in FTF. Both modes, however, give learners the same amount of opportunities to modify output after receiving feedback.

Lai and Zhao's (2006) efforts merit recognition as the first to use a CMC video for the stimulus in SR methodology. However, there are some key limitations in their procedure that may render their findings ungeneralizable to other contexts. First, the SR sessions took place the day after the treatment. The timing of the procedure can be a crucial "procedural pitfall" for SRs (Gass & Mackey, 2000, p. 84), and waiting too long may result in participants not being able to accurately recall what they were thinking, or recalling significantly fewer cognitive events (Garner, 1988, cited in Gass & Mackey, 2000). Second, participants in their study watched the entire FTF video recording but only watched feedback clips of the CMC video recording to prevent participants from becoming "tired, bored, or distracted" (p. 104), given that the CMC videos were longer. Arguably, this difference could have primed learners to be more focused on the purpose of the experiment during the CMC protocols. The third methodological limitation is the questions used by the researchers during the SR sessions. In their coding protocol description, Lai and Zhao exemplify how they carried out the SR procedure:

> Researcher: Ok. Did you notice that he actually corrected your mistake
> Participant D: Ah, I don't, I don't do ah notice. So . . .
> Researcher: . . . but now you noticed, right?
> Participant D: yes, yes
> Researcher: at that time you didn't. You just thought that he agreed with you. (p. 108)

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

These questions and, particularly, the language used by the researchers to elicit online recall may have compromised the SR procedure. In their methodological review on how to conduct protocols, Gass and Mackey (2000) recommend that researchers not pursue "fishing" for answers if participants indicate that they do not remember. Doing so may render an answer that is based on what the participant thinks now and not his or her memory of the event (p. 154). Gass and Mackey also recommend avoiding three-part exchanges during the recall (i.e., acting as a conversational partner) because doing so can alter the nature of the SR (p. 60). In other words, the researcher cannot lead the participant and should prevent participants from analyzing past actions in the present. To this end, researchers must use language to elicit what the learner was thinking at the time of the event itself (i.e., *What were you thinking then?*), focusing clearly on the time of the recall.

To date, no study has used SR to investigate learners' accuracy in noticing types of feedback other than negotiation of meaning and recasts in relation to error types as well as the usability of that feedback (i.e., in terms of opportunities for and production of modified output) in the FTF versus CMC modes. On the basis of Mackey et al.'s (2000) finding that learners were not very accurate at noticing feedback that addressed morphosyntactic errors, more studies are needed to determine if this same result will apply in other language learning contexts. Moreover, given the increase in the development of online language classes, additional research is needed that asserts the generalizability of Mackey et al.'s (2000) and Lai and Zhao's (2006) findings to the CMC environment and that examines whether modality differentially mediates learners' noticing of, opportunity for, and use of feedback. In light of the gaps identified in the literature review, the research questions motivating the current study were as follows:[1]

1. Do learners perceive feedback provided during task-based interaction?
2. Do learners recognize the target of the feedback provided during task-based interaction?
3. Does learner perception of feedback differ according to the mode in which it is provided (i.e., CMC or FTF)?
4. Does opportunity for learner-modified output differ according to mode?
5. Does learner use of opportunity for modified output differ according to the mode in which it is provided?

METHOD

Participants

The participants in this study were 24 intermediate learners of Spanish enrolled in four sections of an intermediate-level Spanish foreign language

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

course at a large university in the United States. The sections followed a task-based, communicative language teaching approach and shared a textbook and syllabus. All learners had experience with information-gap tasks similar to those used in the current study as well as previous experience with the vocabulary and grammar needed to successfully complete the tasks. All were native speakers of English, and some had previous foreign language experience such as Latin, Italian, Greek, French, and Korean. Fourteen participants were females, 10 were males, and their ages ranged from 18 to 21 years. Participants reported being very comfortable using computers, spending an average of 20.5 hr per week on them and more than 2.7 hr per week chatting via various CMC platforms. Participant biodata were collected via a background questionnaire, which is described in the Materials section.

The interlocutor was one of the researchers (henceforth referred to as the "interlocutor"). At the time of the study, the interlocutor was a 28-year-old female doctoral candidate in Spanish applied linguistics who was a near-native speaker of Spanish (native language English) and was trained in task-based language teaching and interaction-based research with 7 years of foreign language teaching experience at both university and elementary levels.

Materials

Materials used in this study included two experimental tasks, SR questions targeting learners' perception of feedback in the two modes, and a questionnaire targeting participant demographics and mode preference.

Experimental Tasks. The experimental information-gap tasks were chosen for this study for two reasons: (a) their similarity to tasks used previously in the students' usual language classrooms and (b) their success in a previous study comparing interaction in CMC and FTF modes with intermediate-level learners at the same university (Baralt & Gurzynski-Weiss, 2011). One task was based on a living room scene and the other was based on a kitchen scene. Each was accompanied by a black-and-white photocopy of the room scene with 12 target vocabulary items highlighted in blue (see Baralt & Gurzynski-Weiss, 2011).

Stimulated Recall Questions. The SR protocols targeted learner perception of feedback during the task-based interaction in both modes. Learners were shown interaction episodes[2] that contained (a) an error, (b) an interlocutor response, and (c) the student response (if any) and were asked what they remembered thinking at the time. Protocols targeting interaction in the CMC mode lasted an average of 10.2 min

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

(range 7.2–14.4); for FTF interaction, protocols lasted an average of 12.5 minutes (range 8.4–18.1). The complete list of questions can be seen in the Appendix.

Demographic Information and Mode Questionnaire. A background questionnaire was administered to obtain information about the participants (including computer and CMC use) and to solicit learners' opinions and perceptions about completing the information-gap tasks in the two modes. Participants were asked to indicate their year of study, major, sex, native language, and other languages studied as well as their motivation for taking Spanish as a foreign language. Additionally, learners were asked how much time they spend using the computer per week, which mode (CMC or FTF) was more enjoyable, which was more challenging, in which they learned more, and which mode they preferred and why. The questionnaire was used both to ensure that learners were comparably familiar with communicating in both FTF and CMC modes and to provide a more qualitative picture of any patterns indicated by the quantitative analyses.

Procedure

Each participant interacted with the interlocutor individually to complete an information-gap task in each of the two modes in a one-shot, experimental design. Both the mode of the task and the task itself were counterbalanced in a Latin square design to prevent carry-over effects from one mode (or task type) to the other. The tasks were completed in Spanish, and the interlocutor provided feedback whenever it seemed natural to do so,[3] addressing errors of lexis, morphosyntax, semantics, and phonology (specific to the FTF mode) or spelling (unique to CMC). The second researcher (henceforth referred to as the "researcher") was in the room during the interactions to record the data needed for the SR protocols.

When completing the task in the CMC mode, the participant and interlocutor sat on opposite sides of a computer lab (reserved for this experiment and otherwise vacant). The information-gap tasks used in this study (from Baralt & Gurzynski-Weiss, 2011) required participants to explain to the interlocutor how to arrange furniture in a room. The participants were given a picture of a room (either a living room or a kitchen) filled with furniture and were instructed to communicate to their partner via iChat[4] the placement of each piece of furniture using Spanish. The participants were told that their partner, the interlocutor, had an empty room and cutouts of the furniture and that the goal of the task was for both pictures to match at the end. While the learner and interlocutor communicated via iChat, the researcher recorded the interaction

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

using iShowU HD, capturing the iChat interaction in real-time video. The recording served as the stimulus for the SRs for the CMC modality.

When completing the task in the FTF mode, the participant and interlocutor sat across from each other at a table with a barrier between them, blocking the task pictures. The participant was given a picture of the second room and instructed to communicate to his or her partner the placement of the furniture items using spoken Spanish; the task goal was again for both rooms to match at the completion of the task. The researcher videotaped the interaction by placing a Sony Handycam camera on a tripod in the back of the computer laboratory; the video of the interaction served as the stimulus for the FTF SRs.

Regardless of mode, immediately after each task was completed, the researcher conducted a SR protocol with the participant while the interlocutor waited outside of the laboratory. The protocols focused on approximately 10 interaction episodes,[5] explained further in the Coding section, with the goal of uncovering each participant's perception of feedback during the task-based interaction.[6] The SR targeting FTF interaction was conducted in front of a TV while the participant and researcher watched clips from the interaction. The researcher paused the tape after each interaction episode and asked the participant what he or she remembered thinking at the time and what he or she believed his or her partner was trying to communicate in that moment (see the questions utilized in the Appendix). The participant also had the opportunity to stop the tape at any time and provide general comments whenever he or she felt it necessary. For the SR targeting CMC interaction, the participant and researcher watched interaction episodes on a computer while the CMC recording served as the stimulus. Stimulated recalls were audiotaped using a digital recorder.

After the first task and mode-specific SR were completed, the learner was offered a break (which no one utilized) and subsequently began the second task, followed by the second SR. The tasks were not timed to ensure that the interaction stayed true to the modes, as research has demonstrated that interaction in the CMC mode takes considerably longer than FTF interaction (e.g., Baralt & Gurzynski-Weiss, 2011; Lai & Zhao, 2006). In this study, time on task ranged from 10 to 37 min in the CMC mode (M = 25.5 min) and from 7 to 25 min in FTF (M = 13 min). All learners successfully completed the tasks, as measured by picture comparison at the end of the task.

Coding and Analysis

The dataset consisted of the FTF and CMC interaction episodes and the SR comments recorded during each protocol. From the interaction episodes,

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

data regarding (a) the nature and amount of student errors, (b) interlocutor feedback, and (c) opportunities for and learner production of modified output were coded. Stimulated recall comments were coded for reported learner perception.

Inter- and Intracoder Reliability. The dataset was coded as follows: First, the researchers transcribed the FTF interaction data and the SR protocols and also tabulated all pertinent interaction data from the CMC chats. Blinding procedures were not necessary as all data were saved with numbers during the original data collection (with identifying information kept separately from the dataset). Each researcher transcribed half of the data and then another 30% of the other researcher's half of the data, selected randomly, to ensure reliability with transcriptions. After the data were transcribed, the researchers created a coding protocol based on Mackey et al. (2000), including the operationalizations and examples from that article. The researchers randomly selected two participants' data of each type (FTF interaction, CMC interaction, FTF SR protocol, and CMC SR protocol) to serve as the pilot data for coding protocols; coding was then discussed between the researchers to ensure consistent categorization. Adjustments were made to the coding protocol on the basis of mutual decision and, in some cases, additional investigation as to the coding norms in recent studies. Examples from the current study replaced the examples from Mackey et al. (2000); the final coding protocol is reported in the next section. The researchers then divided the data equally (and randomly) and each worked on one type of data (e.g., the FTF interaction) at a time. They then recoded the data 2 weeks later, checking the second iteration of coding against their own previous coding to ensure intracoder reliability. After confirming consistent intrarater categorization, the researchers then sent each other 30% of their data for interrater coding. Once all data of a given type were finished, the researchers met again to discuss the coding decisions and to come to an agreement on any discrepancies; the next type of data was subsequently coded in turn. Cohen's kappa intercoder reliability measures were high: for error type, $\kappa = .97$; feedback type, $\kappa = .91$; SR comment type, $\kappa = .95$; opportunity for modified output, $\kappa = .98$; and learner-produced modified output, $\kappa = .96$.

The following sections include operationalizations and original examples from the current Spanish data set as well as English translations underneath. These served as the coding protocol for this study. All errors are highlighted in bold. Capitalization is used to denote the stressed words from the FTF mode; if the item is from the CMC mode, the interaction is represented as it appeared in CMC (e.g., italicized, bolded, etc.).

Error and Feedback Types. Participant errors coded in both modes included lexis, semantics, morphosyntax, and other. Phonological errors

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

were only possible in the FTF mode, and spelling errors were limited to the CMC mode. Thus, for each mode there were five possible error types.

Lexical errors were operationalized as participant use of an incorrect word for the given context or a word that does not exist in the target language. In (1), the participant used a word that does not exist in Spanish, *compactador*, instead of the actual Spanish word, *compresor* "compactor."

(1) Lexis *Hay un* ***compactador*** *de basura.*
"There is a **compactor** of trash."

Errors in semantics were operationalized as errors in communicating meaning, errors in creating understanding, or being unable to express an intended meaning or create understanding. In example (2), the participant was attempting to describe the placement of the recipe book in the kitchen scene but was unable to express his intended meaning to the interlocutor.

(2) Semantics *Hay un libro en . . .* ***el lugar****.*
"There is a book in . . . **the place.**"

A morphosyntactic error was operationalized as dealing with missing words or incorrect word order; lack of gender or number agreement; or incorrect verb conjugation, tense, or mood. In (3), the participant needed to combine the preposition *de* "of" with the article *el* to form *del*, according to the morphosyntactic rules of the Spanish language.

(3) Morphosyntax *Está al lado* ***de el*** *tablón.*
"It's to the side **of the** counter."

Phonological errors in the FTF mode included incorrect accent placement or a mispronunciation violating the rules of the target language (e.g., pronunciation of [h] at the beginning of a Spanish word that was spelled with "h"). In (4), the participant incorrectly placed the accentuation on the first syllable.

(4) Phonology *. . . encima del* ***líb****rero.*
". . . on top of the **book**shelf."

Spelling errors in the CMC mode were errors in the written form of a word. These errors were differentiated from lexical errors because, in these cases, it was clear that the learner had the correct word; the error was simply an issue of spelling that did not change the meaning of a word (which would constitute an error in lexis). Here, in (5), the participant misspelled the word for "machine," *máquina*, but it was clear she understood the correct word to use.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

(5) Spelling . . . *de la* ***máchina*** *para lavar platos.*
". . . of the **machine** to wash dishes."

Finally, a sixth category was reserved for errors that did not fit into the other categories, such as off-target responses. In (6), the participant focused on an object that was part of the room and not an item to be negotiated.

(6) Other . . . *en la ventana* ***es*** *cortinas?*
". . . on the window **is** curtains?"

Interlocutor Feedback. Unlike the earlier studies that focused on one or two types of feedback, in the current study, 10 types of feedback occurred naturally in the data set and were subsequently operationalized and coded for: (a) no feedback, (b) overt correction, (c) recast, (d) metalinguistic feedback, (e) elaboration, (f) elicitation, (g) repetition, (h) confirmation check, (i) clarification request, and (j) negotiation for meaning. The coding *no feedback* was used when the interlocutor did not provide feedback in response to the participant error, as seen in (7), in which, given its preposed position, the correct form of the adjective "big" would be *gran* rather than *grande*.

(7) No feedback
Participant: . . . *al lado del* ***grande*** *sofá.*
Interlocutor: *¿Qué más?*
Participant: ". . . to the side of the **big** sofa."
Interlocutor: "What else?"

An overt correction in the current study, as seen in FTF example (8), was an utterance that provided the participant with the correct form in a way that overtly indicated that the original utterance was incorrect.

(8) Overt correction
Participant: ***Es*** *debajo de—*
Interlocutor: (interrupts) *ESTÁ debajo de . . .*
Participant: "**It is** below**—**"
Interlocutor: "IT IS [different copula] below . . ."

Recasts were utterances that rephrased the learners' original utterances by changing at least one component while keeping the central meaning the same. An example is in (9), in which only the final vowel was changed to reach the targetlike word *derecha* "right":

(9) Recast Participant: *A la derech****o*** *hay una silla.*
Interlocutor: *A la derecha hay una silla, muy bien.*
Participant: "To the **right** there's a chair."
Interlocutor: "To the right there's a chair, very good."

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Feedback that provided the learner with specific information about the nature of their error was coded as *metalinguistic information*, seen in (10). In the example in (10), the student used the adverb *adentro*, which is usually reserved to describe motion, instead of the correct choice, *dentro*, which denotes static placement.

(10) Metalinguistic information

Participant: *Hay un libro y está* ***adentro*** *de la nevera.*
Interlocutor: *Dentro significa que el libro está dentro de la nevera . . .*
Participant: "There's a book and it's **inside of** the refrigerator."
Interlocutor: "Inside means that the book is inside of the refrigerator . . ."

Utterances providing additional semantic information following a participant's expressed confusion were coded as *elaborations*, as seen in (11).

(11) Elaboration

Participant: *¿Libro de recetas por teléfonos?*
Interlocutor: *Las recetas son las que sigues— son como instrucciones para saber cómo se prepara la comida.*
Participant: "Recipe book for phones?"
Interlocutor: "Recipes are what you follow—they're like instructions to know how one prepares the food."

Elicitations were operationalized as attempts to elicit the correct construction, as seen in (12).

(12) Elicitation

Participant: *Al lado está el colgador para papel de* ***toal-las****.*
Interlocutor: *¿Cómo se llama otra vez?*
Student: "To the side is the paper **towel** holder."
Interlocutor: "How is it called again?"

A mutual effort between the interlocutor and participant to understand each other during a series of two or more exchanges was coded as negotiation of meaning, seen in example (13).

(13) Negotiation

Participant: *Um . . . es donde se puede poner fotos, de la famila, um . . . debajo de esto es una . . . um . . . es donde hay una . . . es como está encima del* ***candle.***
Interlocutor: *¿Una vela?*

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Participant: *Sí. Encima de una vela hay una cosa que hace calor.*
Interlocutor: *Ahh . . . ¡fuego!*
Participant: *Fuego, sí.*
Interlocutor: *Donde se hace fuego. ¡La chimenea!*
Participant: *La chimenea. Sí.*
Participant: "Um . . . it's where one can put photos, of the family, um . . . under it is a . . . um . . . is where a . . . is like it is on top of the **candle.**"
Interlocutor: "A candle?"
Participant: "Yes. On top of a candle there's a thing that makes heat."
Interlocutor: "Ohh, fire!"
Participant: "Fire, yes."
Interlocutor: "Where one makes fire! The fireplace!"
Participant: "The fireplace, yes."

Confirmation checks were operationalized as utterances used to ensure that the interlocutor understood what the participant said. An example can be seen in (14).

(14) Confirmation check

Participant: *Hay **un** luz.*
Interlocutor: *¿Una luz?*
Participant: "There's **a** [masculine] light."
Interlocutor: "A [feminine] light?"

Clarification requests were expressions designed to elicit clarifying information with respect to the participant's previous utterance, as seen in (15).

(15) Clarification request

Participant: *En **el otro** pared.*
Interlocutor: *¿Lo pongo dónde exactamente?*
Participant: "On **the other** wall."
Interlocutor: "Where do I put it exactly?"

A repetition was an utterance that repeated and highlighted the learner's error, as in (16). In this example, the student was attempting to describe the placement of the window, which was above the plant. Instead, the student said it was below the plant.

(16) Repetition

Participant: ***La ventana** es debajo de la planta.*
Interlocutor: *¿La ventana?. . . No entiendo.*
Participant: "**The window** is under the plant."
Interlocutor: "The window? I don't understand."

A breakdown of the linguistic content of the feedback episodes can be seen in Table 1.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Table 1. Linguistic content of feedback episodes

	CMC (152 episodes)		FTF (164 episodes)	
Error type	Number	Percentage	Number	Percentage
Morphosyntactic	31	20.4	41	25.0
Phonology/spelling	1	0.7	5	3.0
Semantic	13	8.6	12	7.3
Lexical	107	70.4	106	64.6

Opportunities for and Learner Production of Modified Output. With respect to learner-modified output, the researchers first coded for opportunities for modified output and, once these opportunities had been identified, then coded for subsequent learner-produced modified output.

Opportunities for modified output. When the interlocutor provided time and space for the learner to modify his output in the turn immediately following feedback, this was coded as *opportunity for modified output*. In the CMC example in (17), the interlocutor provided feedback and then stopped typing, which provided the participant time and space to respond.

(17) Opportunity for modified output

Participant: *La* ***refrigiadora.***
Interlocutor: *La NEVERA.*
Participant: *Nevera, sí.*
Participant: "The **refrigerator.**"
Interlocutor: "The REFRIGERATOR."
Participant: "Refrigerator, yes."

In contrast, *no opportunity for modified output* was operationalized as an instance in which the interlocutor did not provide space or time for the learner to modify his or her output in the turn immediately following feedback. In example (18), the interlocutor wrote two lines of text in succession immediately following the feedback, which rendered a modification of output impossible in the discourse by not providing space or time for the participant to respond.

(18) No opportunity for modified output

Participant: ***Es*** *arriba del tablón.*
Interlocutor: *Está arriba del tablón.*
Interlocutor: *Ok*
Interlocutor: *¿Qué más?*
Participant: "**It is** on top of the counter."

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Interlocutor: "It is [different copula] on top of the counter."
Interlocutor: "Ok"
Interlocutor: "What else?"

Learner-modified output. Next, the researchers examined those instances in which the participants did have the opportunity to produce modified output. For instances of *no modified output*, the participant did not modify their output in the turn immediately following interlocutor feedback. In (19), despite the interlocutor's overt correction of the correct lexical item, *está* "is" (from the verb *estar*) instead of *es* "is" (from the verb *ser*), the participant did not modify her output.

(19) No modified output

Participant: ***Es*** *a la derecha de la mesa.*
Interlocutor: *ESTÁ a la derecha de la mesa.*
Participant: *Sí.*
Participant: "**It is** to the right of the table."
Interlocutor: "IT IS [different copula] to the right of the table."
Participant: "Yes."

Instances of modified output were identified when the participant accurately incorporated at least one of the corrections in the turn immediately following feedback. In (20), there are several mistakes: The participant (a) used the incorrect masculine gender for a direction; (b) used the incorrect gender for the word *sofá* "sofa," which is masculine; and (c) did not include the required *a* and article *la* when expressing directions. As seen in example (20), the interlocutor recasted the correct form of all four errors, and the participant correctly modified his output for *a, la*, and the gender of *derecha* but did not modify his output of the gender of *sofá*.

(20) Modified output

Participant: *Está* ***derecho*** *de* ***la*** *sofá.*
Interlocutor: *Está a la derecha del sofá.*
Participant: *Sí, a la derecha.*
Participant: "It's **right** of **the** [feminine] sofa."
Interlocutor: "It's to the right of the [masculine] couch."
Participant: "Yes, to the right."

Stimulated Recall Comments. Each mode had seven possible SR comment categories. Both had lexis, semantics, morphosyntax, mode, no content, and unclassifiable. For the FTF mode, comments also included phonology; for CMC, spelling. Each of these comment categories is examined in turn.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Participants' SR comments coded as *lexis* were operationalized as specific comments about a known or unknown word or a synonym or antonym. An example of an error in lexis is provided in (21).

(21) Lexis "Yeah, I knew that smoking was *fumar* and I didn't remember what flame or fire was so I was trying to go near that and trying to describe that it has a flame."

Comments coded as *semantic* were general comments about communicating meaning or understanding or being unable to express an intended meaning, as seen in (22).

(22) Semantics "I thought I was miscommunicating what I was trying to say. I could tell she didn't understand me."

Morphosyntactic comments were operationalized as targeting sentence formation, structure, or word order or were comments on aspects of grammar such as agreement, tense, or mood, as in (23).

(23) Morphosyntax "I was trying to decide if I needed the subjunctive or not."

Comments were coded as dealing with phonology if they were about pronunciation or accentuation of a certain word or sound in Spanish, as in (24); this category only pertained to the FTF mode.

(24) Phonology ". . . that I was pronouncing it wrong."

In contrast, unique to the CMC mode were comments dealing with spelling a particular word or phrase in Spanish, as seen in (25).

(25) Spelling "I knew what I wanted to write but I wasn't sure how to spell it."

When the participant had nothing to say or did not remember anything about the particular moment, the comment was coded as *no content*, as in (26).

(26) No content "Yeah, I don't remember."

Due to the cross-modal focus of the study and the comments made during the SR protocols, the researchers felt it was prudent to create a category specific to mode. Thus, in (27), the reader can find an example of a *mode* category comment in which a participant commented about the interaction mode (CMC or FTF).

(27) Mode "I liked [chat] because I could look at what I typed."

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Finally, SR comments were coded as *unclassifiable* if the participant made comments that did not fit into one of the aforementioned operationalizations, as in (28).

(28) Unclassifiable "I was frustrated."

The linguistic content of the SR comments in the current study is tabulated in Table 2.

RESULTS

This study sought to investigate differences in feedback perception and utilization during CMC and FTF task-based interaction. Results for each research question are reported in the following sections.

Research Question 1: Learner Feedback Perception in Task-Based Interaction

The first research question asked if learners perceived feedback provided during task-based interaction in FTF and CMC modes. These results can be seen in Table 3. A visual inspection of the data in Table 3 shows that learners did accurately perceive feedback the majority of the time: overall, 68.3% of the time in the FTF mode and 71.1% in CMC. Thus, the answer to the first research question is positive: Learners do accurately perceive feedback as feedback in task-based interaction for well over half of all feedback episodes provided in both modes.

Table 2. Linguistic content of stimulated recall comments

	CMC (152 episodes)		FTF (164 episodes)	
Comment focus	Number	Percentage	Number	Percentage
Lexical	91	59.9	97	59.1
Semantic	21	13.8	24	14.6
No content	17	11.2	19	11.6
Morphosyntactic	15	9.9	18	11.0
Unclassifiable	4	2.6	4	2.4
Mode	4	2.6	0	0.0
Phonological	0	0.0	2	1.2

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Table 3. Learners' accurate perceptions of feedback as measured by stimulated recall protocols

	CMC (152 episodes)		FTF (164 episodes)	
Feedback perceived	Number	Percentage	Number	Percentage
Morphosyntax	15/31	48.4	17/41	41.5
Phonology/spelling	0/1	0.0	2/5	40.0
Semantics	9/13	69.2	8/12	66.7
Lexis	84/107	78.5	85/106	80.2
Total	108/152	71.1	112/164	68.3

Research Question 2: Learner Perception of Feedback Target

The second question asked if learners correctly perceived the target of feedback provided in task-based interaction in FTF and CMC modes. Tables 4–7 illuminate how learners perceived the target of feedback (i.e., lexis, semantics, morphosyntax, phonology/spelling) in each mode. A summary of learners' accuracy of the feedback targets in each mode can be seen in Figure 1.

As shown in Tables 4–7 and in Figure 1, learners were most accurate in their perception of feedback targeting lexis, meaning that they perceived feedback targeting lexis as targeting lexis, according to their SR comments. This occurred 80.2% of the time in the FTF mode and 78.5% of the time in the CMC mode. Learners were the second most accurate in their perception of feedback targeting semantics; they correctly identified the target as semantics 66.7% of the time in the FTF mode and 69.2% of the time in CMC. Next were learners' interpretations of feedback targeting morphosyntax, which they accurately perceived 41.5% of the time in FTF mode and 48.4% of the time in CMC. Finally, for phonology,

Table 4. Learner perceptions of feedback targeting lexis

	CMC (107 episodes)		FTF (106 episodes)	
Comment focus	Number	Percentage	Number	Percentage
Morphosyntax	0	0.0	1	0.9
Phonology/spelling	0	0.0	0	0.0
Semantics	8	7.5	11	10.4
Lexis	84	78.5	85	80.2
No content	11	10.3	7	6.6
Unclassifiable	2	1.9	2	1.9
Mode	2	1.9	0	0.0

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Table 5. Learner perceptions of feedback targeting semantics

	CMC (13 episodes)		FTF (12 episodes)	
Comment focus	Number	Percentage	Number	Percentage
Morphosyntax	0	0.0	0	0.0
Phonology/spelling	0	0.0	0	0.0
Semantics	9	69.2	8	66.7
Lexis	3	23.1	4	33.3
No content	0	0.0	0	0.0
Unclassifiable	0	0.0	0	0.0
Mode	1	7.7	0	0.0

learners accurately perceived feedback targeting phonology 40% of the time in the FTF mode; the CMC counterpart, spelling, was not accurately perceived for the one instance in the dataset.

In sum, for the second research question, learners were accurate in their perception for most of the feedback targets, particularly feedback targeting lexis and semantics, followed by morphosyntax. Learners were least accurate in their perception of feedback targeting phonology (FTF) or spelling (CMC) in the current study.

Research Question 3: Differences in Learner Feedback Perception According to Mode

To answer the third research question, which asked if there were any differences in learner feedback perception in FTF mode as compared to CMC, a series of paired-samples *t* tests were performed; conditions for use of this parametric test were met. The analyses indicated no statistical

Table 6. Learner perceptions of feedback targeting morphosyntax

	CMC (31 episodes)		FTF (41 episodes)	
Comment focus	Number	Percentage	Number	Percentage
Morphosyntax	15	48.4	17	41.5
Phonology/spelling	0	0.0	0	0.0
Semantics	4	12.9	5	12.2
Lexis	4	12.9	6	14.6
No content	6	19.4	11	26.8
Unclassifiable	2	6.5	2	4.9
Mode	0	0.0	0	0.0

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Table 7. Learner perceptions of feedback targeting phonology (FTF) or spelling (CMC)

	CMC (1 episode)		FTF (5 episodes)	
Comment focus	Number	Percentage	Number	Percentage
Morphosyntax	0	0.0	0	0.0
Phonology/spelling	0	0.0	2	40.0
Semantics	0	0.0	0	0.0
Lexis	0	0.0	2	40.0
No content	0	0.0	1	20.0
Unclassifiable	0	0.0	0	0.0
Mode	1	100.0	0	0.0

differences between learners' perception of feedback in FTF versus CMC mode on any of the linguistic foci addressed. There were no statistical differences between how learners perceived morphological feedback in FTF mode (M = 0.67 feedback episodes correctly perceived as targeting morphosyntax, SD = 0.64) compared to CMC (M = 0.63, SD = 0.82), $t(23) = -0.20$, $p = .85$, $d = 0.05$. No statistical differences were found for learner perception of lexical feedback in FTF mode (M = 3.54, SD = 1.93) compared to CMC (M = 3.50, SD = 1.72), $t(23) = -0.08$, $p = .94$, $d = 0.02$; nor for perception of semantic feedback in FTF mode (M = 0.33, SD = 0.64) compared to CMC (M = 0.38, SD = 0.65), $t(23) = 0.23$, $p = .82$, $d = 0.08$. There were also no statistical differences in learners' perception of

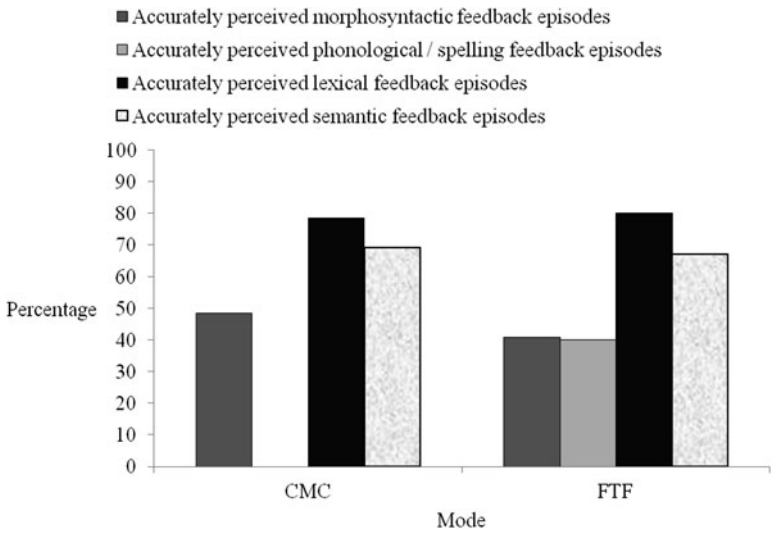


Figure 1. Feedback accuracy according to target and mode.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

feedback addressing phonology (FTF; M = 0.08, SD = 0.28) or spelling (CMC; M = 0.00, SD = 0.00), $t(23)$ = –1.15, p = .16, d = 0.40; however, this finding should be interpreted with caution given that so few instances of phonological and spelling feedback took place (1 and 5 instances in the CMC and FTF modes, respectively, compared to 107 and 106 for lexis, 31 and 41 for morphosyntax, and 13 and 12 for semantics). These results suggest that the mode in which learners interacted did not have a differential effect on how they perceived correctional feedback for any of the linguistic targets addressed by feedback. Hence, the answer to the third research question is negative: Learners did not perceive feedback more accurately in one mode as compared to the other.

Research Question 4: Opportunity for Learner-Modified Output According to Mode

The fourth research question looked at whether there were differences in opportunities for learners to modify their output according to mode. There were many more opportunities for learner-modified output in the FTF mode (i.e., 91.5% of the time) as compared to the CMC mode (i.e., 53.3% of the time). These numbers are further broken down in Table 8.

To see if these differences were statistical, a paired-samples t test was conducted to compare the total number of opportunities for learner-modified output in the FTF mode with those provided in the CMC mode. Once again, conditions for use of this parametric test were met. The analysis revealed a statistical difference in the number of opportunities learners had to produce modified output in FTF communication (M = 6.21, SD = 2.40) versus in CMC (M = 3.38, SD = 2.12), $t(23)$ = 5.03, p < .001, d = 1.25. Modality, therefore, had a significant effect on how many opportunities learners were provided to modify their output.

Table 8. Opportunities for modified output according to mode and error type

	CMC (152 episodes)		FTF (164 episodes)	
Error type	No opp.	Opportunity	No opp.	Opportunity
After morphosyntax error	17	14	6	32
After phonology/spelling error	0	1	1	4
After lexical error	53	54	3	103
After semantic error	1	12	4	11
Total	71/152	81/152	14/164	150/164

Note. No opp. = no opportunity.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Interaction in the FTF mode resulted in statistically more opportunities to modify output than in CMC, and the effect size for this difference was quite large—more than one standard deviation.

Breaking down the analysis further, we then compared opportunities to modify output in each mode according to error type. Paired-samples *t* tests yielded a statistical difference in opportunity to modify output after receiving feedback on lexical errors in FTF mode ($M = 4.29$, $SD = 2.01$) compared to CMC ($M = 2.25$, $SD = 1.60$), $t(23) = 4.22$, $p < .001$, $d = 1.12$, and after feedback addressing morphosyntax in FTF mode ($M = 1.29$, $SD = 1.12$) compared to CMC ($M = 0.58$, $SD = 0.78$), $t(23) = 2.17$, $p = .04$, $d = 0.74$. Differences per mode after feedback addressing phonology or spelling neared significance (FTF: $M = 0.17$, $SD = 0.38$; CMC: $M = 0.04$, $SD = 0.20$), $t(23) = 1.81$, $p = .08$, $d = 0.43$. There were no statistical differences per mode in opportunities for modified output after semantic errors (FTF: $M = 0.46$, $SD = 0.72$; CMC: $M = 0.50$, $SD = 0.78$), $t(23) = -0.21$, $p = .83$, $d = 0.05$. Thus, for this further analysis that compared opportunities for modified output in CMC versus FTF environments for each error type, there were significantly more opportunities for learner-modified output in the FTF mode after feedback that addressed lexical and morphosyntactic errors as well as more opportunities for modified output after phonological errors in FTF mode compared to after spelling errors in CMC, though these differences were not significant. Opportunity to modify output after semantic errors was about the same in both modes.

Research Question 5: Learner-Modified Output According to Mode

A paired-samples *t* test was conducted to examine the fifth research question, which asked whether there were differences in learner-modified output after receiving feedback in FTF compared to CMC mode; assumptions for use of this parametric test were met. The results (see Table 9) were statistically significant, with interaction in the FTF mode resulting in statistically more learner-modified output ($M = 4.21$, $SD = 2.0$) than in the CMC mode ($M = 2.08$, $SD = 1.28$), $t(23) = 6.88$, $p < .001$, $d = 1.27$. This indicates that, after being provided with feedback and when given opportunities to use that feedback in the discourse, participants modified their output significantly more in FTF mode (i.e., 68% of the time) than in CMC (i.e., 62% of the time). Additionally, just as with the findings for opportunities for modified output, the effect size for the difference in learner-modified output in FTF versus CMC modes was large (again, greater than one standard deviation).

We then inspected differences in learner-modified output according to error type. Results from the paired-samples *t* tests showed statistical differences in learner-modified output after receiving feedback on lexical

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Table 9. Learner-produced modified output (MO) according to mode and error type

	CMC (81 episodes)		FTF (150 episodes)	
Error type	No MO	MO	No MO	MO
After morphosyntax error	10	4	8	24
After phonology/spelling error	1	0	0	4
After lexical error	16	38	36	67
After semantic error	4	8	4	7
Total	31/81	50/81	48/150	102/150

errors in FTF mode (M = 2.79, SD = 1.61) compared to CMC (M = 1.58, SD = 1.18), $t(23) = 3.61$, $p = .001$, $d = 0.86$, and after feedback addressing morphosyntax in FTF mode (M = 0.96, SD = 0.95) compared to CMC mode (M = 0.17, SD = 0.38), $t(23) = 4.16$, $p < .001$, $d = 1.1$. There were also statistical differences in modified output following feedback that addressed phonology or spelling (FTF: M = 0.17, SD = 0.38; CMC: M = 0.00, SD = 0.00), $t(23) = 2.15$, $p = .04$, $d = 0.43$. Again, however, these results should be interpreted with caution due to the paucity of feedback focusing on phonology or spelling present in the data set. Differences per mode in modified output after semantic errors were not statistical (FTF: M = 0.29, SD = 0.55; CMC: M = 0.33, SD = 0.70), $t(23) = -0.25$, $p = .80$, $d = 0.06$. Overall, learners modified their output after feedback targeting errors in lexis, morphosyntax, and phonology or spelling significantly more in FTF mode than in CMC. Modified output after feedback addressing semantics, however, was comparable in both CMC and FTF modes.

Results Summary

In sum, the current study found that participants did perceive feedback in task-based interaction and that they perceived feedback addressing lexis and semantics accurately most of the time, and feedback addressing morphosyntax accurately almost half of the time. They were not as accurate for phonological feedback. Additionally, it was found that participants did not perceive feedback statistically more in one mode over the other. Results indicated that there were, however, significant differences according to mode with respect both to the number of opportunities learners had to modify their output and to the frequency with which participants took advantage of these opportunities to produce modified output.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

DISCUSSION

In the current study, Spanish as a foreign language learners were most accurate in their perception of feedback targeting lexis, much like the IFL and ESL learners in Mackey et al. (2000). However, that is the only finding from Mackey et al.'s study replicated in the current data set. Whereas Mackey et al. reported that learners were second most accurate about feedback targeting phonology, in the present study, learners were second most accurate about feedback targeting semantics (a category that Mackey et al. did not analyze, given that so few instances of feedback targeting semantics took place in their dataset). Another key difference is how accurate learners were at noticing feedback that targeted morphosyntactic errors: Learners in the present study accurately noticed morphosyntactic feedback much more than Mackey et al.'s learners (41.5% in FTF mode and 48.4% in CMC in the current study, compared to Mackey et al.'s result of only 13% for ESL learners and 24% for IFL learners). To investigate if the type of corrective feedback caused this difference, a post hoc analysis was carried out to examine what kind and how much feedback was used to address different error types. A distribution of all of the naturally occurring feedback types and the linguistic errors they targeted are provided in Tables 10 and 11.

Recall that Mackey et al. (2000) reported that among recasts, negotiation of meaning, or a combination of the two, recasts were overwhelmingly the most common type of feedback used to address morphosyntactic errors (i.e., 75%). In the current study, recasts were also the most common type of feedback used to address morphosyntactic errors in the FTF mode. Table 10 shows that recasts, followed by overt corrections, were the most common type of feedback targeting morphosyntactic errors. In CMC, the opposite was observed, with overt corrections being the most common feedback type used to correct morphosyntactic errors, followed by recasts (see Table 11). This finding also shows a key difference in how much feedback was allocated and, specifically, which error types were targeted by the majority of feedback. Recall that nearly half of all of the feedback episodes addressed phonological errors (e.g., erroneous pronunciation) in the ESL learner dataset in Mackey et al.'s study; 18% of feedback targeted phonology in their IFL dataset. In the current study, only five episodes of feedback that targeted pronunciation were given to learners in the FTF mode (i.e., less than 5% of all feedback).

The learners in Mackey et al.'s (2000) study noticed feedback targeting phonological errors much more than the learners in the current study—it was the second most accurately perceived target of feedback in Mackey et al. (2000). Nevertheless, this arguably might have been because they were also provided with much more feedback that addressed nontargetlike pronunciation. Likewise, one could argue that

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Table 10. Distribution of feedback and targeted error type in FTF

Feedback type	Morphosyntax	Phonology/ spelling	Lexis	Semantics
Nothing (1)	1/41 (2.4%)	0	0	0
Overt correction (32)	12/41 (29.3%)	2/5 (40%)	17/106 (16.0%)	1/12 (8.3%)
Recast (56)	23/41 (56.1%)	2/5 (40%)	31/106 (29.2%)	0
Metalinguistic information (1)	1/41 (2.4%)	0	0	0
Elaboration (13)	0	0	9/106 (8.5%)	4/12 (33.3%)
Elicitation (7)	0	1/5 (20%)	4/106 (3.8%)	2/12 (16.7%)
Negotiation for meaning (42)	0	0	38/106 (35.8%)	4/12 (33.3%)
Confirmation check (2)	1/41 (2.4%)	0	1/106 (0.9%)	0
Clarification request (9)	3/41 (7.3%)	0	6/106 (5.7%)	0
Repetition (1)	0	0	0	1/12 (8.3%)
Total (164)	41	5	106	12

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Table 11. Distribution of feedback and targeted error type in CMC

Feedback type	Morphosyntax	Phonology/spelling	Lexis	Semantics
Nothing (0)	0	0	0	0
Overt correction (42)	16/31 (51.6%)	0	25/107 (23.4%)	1/13 (7.7%)
Recast (40)	12/31 (38.7%)	0	28/107 (26.2%)	0
Metalinguistic information (0)	0	0	0	0
Elaboration (11)	1/31 (3.2%)	0	7/107 (6.5%)	3/13 (23.1%)
Elicitation (2)	1/31 (3.2%)	0	1/107 (0.9%)	0
Negotiation for meaning (35)	0	0	29/107 (27.1%)	6/13 (46.2%)
Confirmation check (5)	0	0	5/107 (4.7%)	0
Clarification request (17)	1/31 (3.2%)	1/1 (100%)	12/107 (11.2%)	3/13 (23.1%)
Repetition (0)	0	0	0	0
Total (152)	31	1	107	13

the learners in the current study noticed morphosyntactic feedback more because it was the second highest target of feedback episodes (after lexical errors). In speculating about the cause for dissimilarities among the most common feedback targets during interaction in these two studies, we refer to one key element of interaction research that is often omitted when analyzing corrective feedback: the provider of the feedback. What errors are corrected, how they are corrected, and whether or not learners are provided with opportunities to modify their output after feedback are all concurrent decisions made by the interlocutor. In Mackey et al., the interlocutors were a native speaker of English and a nonnative speaker of Italian. Although no additional details were provided on these interlocutors, we know the following about the interlocutor from the current study: She was a trained SLA researcher focusing on task-based interaction and feedback as part of her research agenda. Though there is still very little research on the role of interlocutor individual differences in L2 acquisition processes, a small number of studies have demonstrated that for instructor interlocutors, factors such as their native language, their educational background (particularly in terms of SLA education), and their years of teaching experience can significantly affect the amount and type of feedback provided as well as the reasoning behind why certain errors are or are not corrected (e.g., Gurzynski-Weiss, 2010). Given that both Mackey et al. and the current study used tasks in which meaning was the central focus (i.e., spot-the-difference and information-gap tasks), we interpret these results to suggest that one reason for which differences between the two studies were observed in the amount and target of feedback provided may very well be explained by the provider of the feedback.

In the present study, there was no statistical difference in learners' accurate perception of feedback that targeted any of the different linguistic foci in CMC mode compared to FTF. Contrary to theoretical and empirical claims that CMC may enhance learner noticing of feedback, our results did not show any significant difference between the different modes in learner noticing of form. An examination of Tables 10 and 11 also does not indicate a major difference in types of feedback provided in the two modes. This indicates that learners can focus on meaning in CMC just as much as they do in FTF communication, especially when the goal of the task is one of meaning negotiation. In interaction-based research, CMC has often been considered an environment with a heightened capacity to direct learners' selective attention to form, given the extra online processing time that learners have during a task. However, it is possible that the oft-cited "extra time" in CMC (e.g., Baralt & Gurzynski-Weiss, 2011; Lai & Zhao, 2006) is mostly the result of the dual-mode requirement to type, read, and process incoming messages all at the same time. Baralt and Gurzynski-Weiss (2011) reported greater average time on task in CMC mode compared to FTF, and they added

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

that this difference was not because more feedback was provided in CMC. Rather, it was because task completion in CMC takes longer. We reason that this may have been the case here.

One key modal difference in the present study was the difference in opportunities for modified output. Learners in the present study had significantly more opportunities to modify their output after feedback addressing lexis, morphosyntax, and phonology in FTF mode than in CMC, and they took advantage of these opportunities significantly more in the FTF mode. This runs counter to the findings of Lai and Zhao (2006), who reported that learners had nearly identical opportunities to modify output after feedback provision in CMC and FTF modes (70% and 71%, respectively). We once again hypothesize that the role of the interlocutor may have mediated the differences between these two studies. In Lai and Zhao's investigation, learners were paired with other learners. Learners in the present study were paired with an expert interlocutor. It might very well have been the case that in the CMC mode, learners in Lai and Zhao's study were not as confident in taking on an expert role and, instead, chose to wait for their partner's response (thereby giving him or her an opportunity to respond to and use the feedback) as opposed to moving forward with the next message. More research is clearly needed on the role of the interlocutor in the provision of opportunities for modified output after feedback.

It is also interesting to note that, despite the feedback differences observed, there were no significant differences across mode in opportunity for modified output and modified output after feedback that specifically addressed semantics. We posit that perhaps when there is an error about meaning (as opposed to form), both the learner and the interlocutor are, by necessity, vested in the exchange to achieve a mutual comprehension. Whereas feedback that addresses errors in lexis, grammar, and pronunciation may or may not result in opportunities for modified output as well as actual modified output, feedback that addresses meaning obligates a modification, and this, in and of itself, will necessitate an opportunity to do so as requisite. Thus, at least in the current study, only when there was a focus on form, rather than on meaning, were clear mode differences found with respect to opportunities for and learner production of modified output.

CONCLUSIONS

The current study demonstrated that Spanish as a foreign language learners, for the most part, perceive corrective feedback provided in task-based interaction with an expert interlocutor in both FTF communication and CMC. Furthermore, and in line with previous literature on the topic, results demonstrated that learners' accuracy with respect to

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

the target of feedback differed according to the type of error being addressed. Although there were differences with previous literature (e.g., Mackey et al., 2000) with regard to which feedback targets were most accurately perceived by learners, we speculate that this is related to how much feedback was provided that addressed specific error targets. Additionally, this study found that the FTF mode gives learners significantly more opportunities to modify their output after receiving feedback and that learners take advantage of these opportunities more in the FTF mode than in the CMC mode.

There are many areas for future research, as, to the best of our knowledge, this is the first study to investigate the accuracy of learner perception of feedback in relation to different feedback types in CMC and FTF environments. In this study, 10 naturally occurring feedback types were accounted for. Whether or not the specific nature of feedback also modulates the provision of opportunities for modified output and learner provision of modified output is clearly an area for future research. For example, it would be interesting to examine if learners are more inclined to modify their output after specific types of feedback over others. Future research would also benefit from qualitatively examining different types of modified output, such as full versus partial output (see Baralt, 2010), and how this relates to perception of feedback.

This study took a first step in researching how modality differentially mediates accurate noticing of feedback and the usability of that feedback. More research is needed that explores the relationship between different types of feedback, learner accuracy about the targets of feedback, and learning outcomes in FTF versus CMC modes. On the basis of our findings, we would argue that future studies must also examine interlocutor or instructor reasoning behind their provision of feedback. So far, studies utilizing SR to explore how accurate learners are in noticing feedback and the target of that feedback have found different results in terms of which errors are most addressed by feedback (e.g., Gass & Lewis's 2007 work with heritage learners; Lai & Zhao, 2006; Mackey et al., 2000); however, these studies have also only looked at two types of feedback. Investigating why certain errors are addressed more than others is another area for future research, as findings have shown that interlocutors differ in their feedback perception and provision according to interlocutor- and learner-related variables (e.g., Gurzynski-Weiss, 2010, in press).

As with all experiments in SLA, this study is not without limitations. First are the inherent limitations with retrospective methodology—namely, that participants will experience memory deterioration once they have finished a task. Although the use of a stimulus is argued to elicit more accurate recall, the accessibility of learners' mental processes is, admittedly, not an undisputed issue. The second limitation is the fact that all CMC interaction was executed in the same lab. Arguably,

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

the computerized element of this study would have been more authentic to online language learning if the interlocutor and participant interactions had been in different physical spaces. This feature was unfortunately unavoidable due to spacing availability during the experiment. Another shortcoming is the fact that this study did not take into account learner-internal cognitive variables that have been shown to mediate the efficacy of corrective feedback, such as working memory capacity. Doing so could have provided a better picture of how and for whom feedback was most effective, especially given that different target structures were investigated in this study. Last but not least, this study did not examine L2 development. Although there are theoretical claims that learners must notice feedback to be able to learn from it, the present study was limited to examinations of whether learners accurately perceived feedback as feedback as well as the usability of feedback in task-based interaction. So far, almost all empirical studies have explored learners' perception of corrective feedback separately from learning as a result of corrective feedback. Follow-up studies that explore both feedback perception and learning concurrently are certainly warranted.

We would like to conclude with this study's plausible pedagogical implications for teachers. First, in task-based one-on-one interaction, learners can and do notice feedback in both FTF and computerized interaction environments. Computer-mediated discourse may, however, result in fewer opportunities for modified output after providing learners with feedback. In his guidelines for the provision of corrective feedback, Ellis (2009) said, "Teachers need to create space following the corrective move for learners to uptake the correction" (p. 14). On the basis of our findings, this guideline may be especially critical for feedback provided during computerized interaction. Additionally, learners perceive feedback differently according to the target of the feedback (i.e., the error that the feedback is correcting). Instructors should take care to provide feedback that is most in line with the goals of a particular task or lesson, as there is evidence that learners pay attention to and work with feedback targeting the most communicatively salient aspects of the interaction and, arguably, the focus of the task.

Received 23 January 2012
Accepted 9 May 2012
Final Version Received 9 September 2012

NOTES

1. The first two research questions are paraphrased from Mackey et al. (2000).
2. These are operationalized with examples in the Coding section.
3. This followed protocol from Mackey et al. (2000). Originally there was concern that, in this study, differences in time on task might have led to a greater amount of

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

negotiation of meaning in CMC as compared to FTF modes. On additional analysis of the data, however, it was found that increased time on task in CMC did not lead to disproportionally increased interaction nor to more feedback provision but was simply the nature of the mode. The increased time in CMC was used by learners to prepare what they would type, review their output prior to sending it, and scroll up to review and utilize previous dialogue, according to comments made in the SR protocols.

4. iChat is an instant messaging application; only written text messages were used for this study.

5. The researcher asked each participant about an average of 10 interaction episodes (ranging from 9 to 15) per SR. Because this was the first study to utilize a chat recording for a CMC stimulus, some of the episodes focused on the uniqueness of the modes (e.g., asking learners what they remember thinking when they were scrolling back up in CMC or gesturing in FTF communication). The time spent on protocols was comparable between modes: an average of 10.2 min for CMC and 12.5 min for FTF.

6. Each participant was also asked about one to two items not dealing with an error. For these items, the researcher focused on mode-specific behavior (e.g., gestures in FTF mode, scrolling back up in CMC to refer to a previous word) to further investigate how interaction differed according to mode.

REFERENCES

Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, *28*, 543–574.

Baralt, M. (2010). *Task complexity, the Cognition Hypothesis, and interaction in CMC and FTF environments* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Baralt, M. (in press). The impact of cognitive complexity on feedback efficacy during online versus face-to-face interactive tasks. *Studies in Second Language Acquisition*.

Baralt, M., & Gurzynski-Weiss, L. (2011). Comparing learners' state anxiety during task-based interaction in computer-mediated and face-to-face communication. *Language Teaching Research*, *15*, 201–229.

Blake, R. J. (2011). Current trends in online language learning. *Annual Review of Applied Linguistics*, *31*, 19–35.

Carpenter, H., Jeon, S., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition*, *28*, 209–236.

Chapelle, C. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. New York: Cambridge University Press.

Egi, T. (2010). Uptake, modified output, and learner perceptions of recasts: Learner responses as language awareness. *Modern Language Journal*, *94*, 1–21.

Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, *1*, 3–18.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Gass, S. M., & Lewis, K. (2007). Perceptions about interactional feedback: Differences between heritage language learners and non-heritage language learners. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 79–99). Oxford: Oxford University Press.

Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Erlbaum.

Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA Review*, *19*, 3–17.

Goo, J. (2012). Feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, *34*, 445–474.

Gurzynski-Weiss, L. (2010). *Factors influencing oral corrective feedback provision in the Spanish foreign language classroom: Investigating instructor native/nonnative speaker status, second language acquisition education, and teaching experience* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Gurzynski-Weiss, L. (in press). Instructor characteristics and classroom-based second language acquisition of Spanish. In K. L. Geeslin (Ed.), *The handbook of second language spanish*. Hoboken, NJ: Wiley Blackwell.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Gurzynski-Weiss, L., & Révész, A. (2012). Tasks, teacher feedback, and modified output in naturally occurring classroom interaction. *Language Learning*, *62*, 851–879.
Havranek, G. (2002). When is corrective feedback most likely to succeed? *International Journal of Educational Research*, *37*, 255–270.
Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, *25*, 1– 36.
Jeon, K. S. (2007). Interaction-driven L2 learning: Characterizing linguistic development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 379–403). Oxford: Oxford University Press.
Keck, C., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Amsterdam: Benjamins.
Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning & Technology*, *10*, 102–120.
Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*, 309–365.
Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, *27*, 361–386.
Loewen, S., & Philp, J. (2006). Recasts in the adult L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal*, *90*, 536–556.
Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition: Vol. 2. Second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Erlbaum.
Lyster, R. (2001). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, *51*, 265–301.
Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *26*, 399–432.
Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*, 37–66.
Lyster, R., & Saito, K. (2010). Interactional feedback as instructional input: A synthesis of classroom SLA research. *Language, Interaction and Acquisition*, *1*, 276–297.
Mackey, A., Gass, S. M., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, *22*, 471–497.
Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 408–452). Oxford: Oxford University Press.
Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, *82*, 338–356.
Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–208). Amsterdam: Benjamins.
McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, *56*, 693–720.
Nassaji, H. (2007). Elicitation and reformulation and their relationship with learner repair in dyadic interaction. *Language Learning*, *57*, 511–548.
Ortega, L. (2009). Interaction and attention to form in L2 text-based computer-mediated communication. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction* (pp. 226–253). New York: Taylor & Francis.
Philp, J. (2003). Constraints on "noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, *25*, 99–126.
Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, *31*, 437–470.
Révész, A., & Han, Z. (2006). Task content familiarity, task type, and efficacy of recasts. *Language Awareness*, *3*, 160–179.
Révész, A., Sachs, R., & Mackey, A. (2011). Task complexity, uptake of recasts, and second language development. In P. Robinson (Ed.), *Researching task complexity: Task demands, task-based language learning and performance* (pp. 203–236). Amsterdam: Benjamins.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 131–164). Amsterdam: Benjamins.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, *13*, 96–120.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu: University of Hawai'i Press.

Sheen, Y. (2007). The effect of corrective feedback, language aptitude and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 301–322). Oxford: Oxford University Press.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.

Swain, M., & Lapkin, S. (2002).Talking it through: Two French immersion learners' response to reformulated writing. *International Review of Applied Linguistics*, *37*, 285–304.

Tarone, E., & Bigelow, M. (2007). Alphabetic print literacy and processing of oral corrective feedback in L2 interaction. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 101–121). Oxford: Oxford University Press.

Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 171–195). Oxford: Oxford University Press.

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press

APPENDIX

STIMULATED RECALL QUESTIONS AND PROTOCOL (AS SEEN AND FOLLOWED BY THE RESEARCHERS DURING THE EXPERIMENT)

Instructions (to Be Read to the Participant at the Beginning of the Session):

What we're going to do now is watch a video of the task you just completed. We are interested in knowing what you were thinking at that time during the task. We can see what you were doing in the video, but we don't know what you were thinking. So every time I stop the tape, I'd like you to tell me what you were thinking at that time when you were talking with your partner. If there is any time that you would like to stop the tape and comment on something, just let me know, okay? Do you have any questions? Ready to begin?

For Feedback Episodes and Other Points That We'd Like the Participant to Comment On:

1. What do you remember thinking at this point during the task?
2. What did you think about your partner at this point?
3. What did you think about your participation in the task at this point?
4. What did you think your partner was trying to communicate to you at this time?
5. Is there anything else you would like to comment on about this moment?

At the End of the Session:

1. What did you think of the task?
2. What did you think about interacting with your partner during the task?
3. What were the easiest things about the task?
4. What were the most difficult things about the task?
5. Is there anything else you would like to add about the task?

https://doi.org/10.1017/S0272263113000363 Published online by Cambridge University Press