CrossMark

Studies in Second Language Acquisition, 2017, **39**, 225–256.
doi:10.1017/S0272263116000449

PRONUNCIATION IN FACE-TO-FACE AND AUDIO-ONLY SYNCHRONOUS COMPUTER-MEDIATED LEARNER INTERACTIONS

Shawn Loewen
Michigan State University

Daniel R. Isbell
Michigan State University

Studies of learner-learner interactions have reported varying degrees of pronunciation-focused discourse, ranging from 1% (Bowles, Toth, & Adams, 2014) to 40% (Bueno-Alastuey, 2013). Including first language (L1) background, modality, and task as variables, this study investigates the role of pronunciation in learner-learner interactions. Thirty English learners in same-L1 or different-L1 dyads were assigned to one of two modes (face-to-face or audio-only synchronous computer-mediated communication) and completed three tasks (picture differences, consensus, conversation). Interactions were coded for language-related episodes (LREs), with 14% focused on pronunciation. Segmental features comprised the majority of pronunciation LREs (90%). Pronunciation LREs were proportionally similar for same-L1 and different-L1 dyads, and communication modality yielded no difference in frequency of pronunciation focus. The consensus task, which included substantial linguistic input, yielded greater pronunciation focus, although the

We would like to thank Dan Reed and Dennie Hoopingarner (Co-PIs) for funding the study's data collection through their International Research and Studies Program Grant: *Articulated Research: Web 2.0 Technology and Instructed Language Acquisition*. In addition, we are indebted to Dominik Wolff and Xuehong (Stella) He for their assistance in data transcription and coding. Finally, we would like to thank the special issue editors for inviting us to be part of this special issue, as well as the anonymous reviewers who provided invaluable feedback.

Correspondence concerning this article should be addressed to Shawn Loewen, Second Language Studies, B255 Wells Hall, 619 Red Cedar Road, Michigan State University, East Lansing, MI 48824-1027. E-mail: loewens@msu.edu

© Cambridge University Press 2017

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

results did not achieve statistical significance. These results help clarify the role of pronunciation in learner-learner interactions and highlight the influence of task features.

It is generally accepted that grammar and, to a lesser degree, vocabulary have been the primary focus of second language acquisition (SLA) and instructed SLA (ISLA) research (Derwing & Munro, 2005; Nassaji & Fotos, 2011). However, although pronunciation does not receive commensurate attention in ISLA, it has not been completely overlooked; indeed, interest in pronunciation teaching and research "is on the upswing" (Levis, 2016, p. 423). For example, Lee, Jang, and Plonsky's (2014) recent meta-analysis of pronunciation instruction included 89 studies, most of them conducted within the previous 10 years. Encouragingly, Lee et al. found a healthy effect size for pronunciation instruction (d = .80 in studies with a control group). Despite this surge of research interest, however, the status of pronunciation instruction in the average second language (L2) classroom may be less rosy: Instructional time devoted to pronunciation is generally minimal (about 6% of class time in Foote, Holtby, & Derwing, 2011), teachers often have low confidence in teaching pronunciation (Derwing & Munro, 2015), and pronunciation is inconsistently addressed by commonly used instructional materials (e.g., English as a L2 [ESL] general-skills textbooks; Derwing, Diepenbroek, & Foote, 2012).

Although the growing number of pronunciation instruction studies have contributed to ISLA knowledge, these primarily represent interventionist approaches in which researchers manipulate specific variables; however, fewer observational studies have been conducted, with the result that less is known about how pronunciation might be topicalized spontaneously during communicative interaction in ISLA contexts. The control provided by interventionist studies can facilitate the assessment of the effectiveness of specific interventions; nevertheless, observational studies afford greater ecological validity and can be used to ensure that interventionist research is relevant to teachers' and students' classroom practices (Loewen & Philp, 2012).

Interventionist research on pronunciation-focused, incidental corrective feedback from teachers (e.g., Saito & Lyster, 2012) has shed some light on the issue, but less is known about the role of pronunciation in learner-learner interaction that is often prevalent in communicative, meaning-oriented classrooms (Pica, 1987; Sato & Ballinger, 2016). Such learner-learner interactions are believed to be beneficial to language development (Adams, 2007; Philp, Adams, & Iwashita, 2014; Sato & Ballinger, 2016), and there is little reason to suggest that pronunciation would be an exception. Furthermore, the continued progress and propagation of

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

communication technology is rapidly altering the landscape of L2 instructional contexts. Nevertheless, computer-assisted language learning, which is now facilitated by interactions over audio-only and audiovisual synchronous computer-mediated communication (SCMC), represents a context in which pronunciation has yet to receive much attention, even though the interaction approach (Gass & Mackey, 2015) would posit that SCMC is another useful context in which L2 interaction and acquisition can occur (Sauro, 2011; Ziegler, 2016a, 2016b).

In communicative tasks, pronunciation that interferes with meaning can lead to a breakdown of communication, requiring negotiation for meaning to take place (Long, 1983, 1996). During this negotiation, L2 learners and their interlocutors work to repair the communicative breakdown and, in doing so, focus on the linguistic trigger of the problem. Often in interaction research, negotiation for meaning has been operationalized as language-related episodes (LREs), which are segments of discourse in which interlocutors topicalize linguistic items, either due to a communication breakdown or a desire or need to focus on linguistic accuracy (Swain & Lapkin, 1998). This brief attention to linguistic form during larger meaning-focused interaction has been shown to be beneficial for L2 development (Mackey & Goo, 2007). However, the role of pronunciation in meaning negotiation has seldom been the focus of learner-learner interaction research (though see Kennedy, Guénette, Murphy, & Allard, 2015, for a recent example in L2 French). Furthermore, there has been little research into factors, such as first language (L1) background or task characteristics, that may mediate attention to pronunciation during learner-learner interaction.

In nonpronunciation-focused studies, interacting with L1-different interlocutors has been found to result in more L2 production (Sato & Fujii, 2013) and to improve learners' oral L2 communication skills (Bueno-Alastuey, 2011). Although to our knowledge no studies have investigated the effects of L1 background pairing on pronunciation during interaction, it is an important issue to consider due to the influence of pronunciation on intelligibility (Derwing & Munro, 2015; Levis, 2016). In terms of tasks in nonpronunciation-focused studies, various task characteristics, such as the type of information exchange (one way vs. two way) and the presence of task essential language, have been investigated, and in general task-based interaction has been found to be beneficial for L2 acquisition (e.g., Keck, Iberri-Shea, Tracy-Venture, & Wa-Mbaleka, 2006). But again, the role of pronunciation, particularly in respect to task characteristics, has received minimal research attention. For example, Keck et al.'s (2006) meta-analysis of task-based interaction considered only lexical and grammatical acquisition; pronunciation was absent. Nevertheless, at least two studies (Saito, 2015; Sicola, 2008) have considered tasks and pronunciation, and these will be reviewed in the following text. In sum, an understanding of how

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

frequently pronunciation issues might be attended to in interaction and how learners navigate pronunciation-related communication failures in common instructional conditions (e.g., face-to-face [FTF] or SCMC, same- or different-L1 pairings, varied task conditions) would provide useful information for ISLA researchers and practitioners. Thus, the present study investigates the role of pronunciation in task-based learner interaction across instructional modes, linguistic backgrounds, and task conditions.

LITERATURE REVIEW

Intelligibility and Comprehensibility in L2 Pronunciation

In pronunciation research, intelligibility has been defined as the "extent to which listeners' perceptions match speakers' intentions (actual understanding)" (Munro & Derwing, 2015, p. 15), and, for communication to occur successfully, a speaker's L2 production must be intelligible for listeners to understand their interlocutors' messages. In interaction, a lapse in intelligibility may occur for any number of reasons, and one cause for such breakdowns is unintelligible sounds in speech. In fact, Levis (2016) claims that "pronunciation is always critical" for intelligibility (p. 424). Further, as Field (2005) notes, intelligibility can be considered "the extent to which the *acoustic-phonetic content* of the message is recognizable by a listener" (p. 401, emphasis added). In this regard, Levis's (2005) Intelligibility Principle reflects the prime importance of speech being understood rather than being nativelike because simply having a foreign accent does not preclude being easily understood by listeners (Munro & Derwing, 1995). This perspective has become prevalent in L2 pronunciation pedagogy and research (Derwing & Munro, 2015). In any discussion of intelligibility, it is important to note the role of the listener: Listener perception is the measure of successful speaker productions. For example, L2 pronunciation researchers examining intelligibility in the context of monologic production have commonly employed listener transcriptions, answers to comprehension questions, or identification tasks to determine whether (or how well) a speaker's message was conveyed (Munro & Derwing, 2015). From the perspective of classroom interaction, it is likely that the listener would signal, in some way, to the speaker when intelligibility has dropped below a critical threshold and understanding has lapsed. Thus, the comprehension skills of the listener and the pronunciation skills of the speaker are both critical in the achievement of communicative success.

Another construct related to intelligibility is *comprehensibility*, or the "perceived degree of difficulty experienced by the listener in understanding speech" (Munro & Derwing, 2015, p. 14). Although there is

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

overlap between these two constructs, comprehensibility emphasizes listener effort, while intelligibility emphasizes successful communication and understanding. Like intelligibility, comprehensibility is measured through listener judgments, but instead of examining the veracity or accuracy of reception, listeners typically provide a subjective rating of their difficulty in comprehending the speaker's production (conventionally on a 9-point Likert scale; Isaacs & Thomson, 2012). Comprehensibility has been found to be influenced by pronunciation as well as several nonpronunciation linguistic factors, such as lexis or grammar (Crowther, Trofimovich, Saito, & Isaacs, 2015; Saito, Trofimovich, & Isaacs, 2016).

An important question regarding these related constructs is how much they are affected by the various aspects of pronunciation. When considering subsets of pronunciation features, a distinction is commonly made between segmental features, consisting of individual or combined consonants and vowels, and suprasegmental features, which go beyond an individual sound and include stress, syllable structure, rhythm, and intonation (Kennedy et al., 2015; Zielinski, 2015). Directly comparing the relative contributions of segmental and suprasegmental features to intelligibility remains a challenge for researchers, possibly due to the wide range of methods used (Munro & Derwing, 2015). However, research investigating comprehensibility frequently examines segmental and suprasegmental features simultaneously and researchers have adopted relatively uniform methods for measuring the construct (i.e., scale-based subjective listener ratings). A body of research involving noninteractive speech (e.g., read-alouds, monologues) has shown that comprehensibility is influenced considerably by both segmental and suprasegmental features of pronunciation (Derwing, Munro, & Wiebe, 1998; Kang, Rubin, & Pickering, 2010; Saito, Trofimovich, & Isaacs, 2015, 2016).

Although research on comprehensibility and its linguistic underpinnings has yielded many insights about L2 speech with some relevance to the present study, the construct is difficult to reconcile in the context of interaction. In interactionist research, rather than examining an interlocutor's perceived degree of comprehension difficulty, evidence of linguistic difficulties is examined in the interaction, where particular linguistic features can trigger an observable breakdown in understanding leading to negotiation for meaning. Consequently, the construct of intelligibility, concerned with actual understanding, is more relevant to interaction. Indeed, interaction research frequently categorizes the linguistic source (i.e., trigger) of the communication breakdown according to grammar, vocabulary, or pronunciation (e.g., Ellis, Basturkmen, & Loewen, 2001; Zhao & Bitchener, 2007). Often in such studies, pronunciation is characterized as consisting of segmental or suprasegmental components.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

The role of intelligibility in interaction is exemplified in several studies, with examples of both segmentally and suprasegmentally driven LREs attested to; however, Jenkins (2002) suggests that most intelligibility breakdowns are related to segmental features. For example, Matsumoto (2011), examining intelligibility in naturalistic interactions among six English-as-a-lingua-franca users attending an American university, described how participants conversing at dinner negotiated meaning when communication broke down due to accent or mispronunciation. Matsumoto's participants often faced intelligibility difficulties due to segmental pronunciation issues (e.g., deletion or substitution of phonemes) but could overcome these problems through accommodation strategies (e.g., requests for clarification, repetition by the speaker or hearer), which allowed for repair. In the context of language tandems (i.e., native speaker [NS]-nonnative speaker [NNS] interactions), Horgues and Scheuer (2014) observed a read-aloud task where English NSs were asked to give feedback to French learners of English. Most corrective feedback given to learners was related to segmentals (95.4%). Although the NSs in the tandems admitted that not all pronunciation features were unintelligible, the researchers identified numerous instances in which mispronunciation of segmental features led to lexical confusion or unintelligibility (such as *hunger* being rendered closer to *anger* by a French learner of English). While Matsumoto (2011) and Horgues and Scheuer (2014) reported primarily segmental pronunciation issues in interaction, other research has attested to suprasegmental features leading to misunderstandings. In Pickering's (2009) study of EFL task-based interactions, sentence stress, pitch movement, and relative pitch were shown to affect intelligibility (with misunderstandings potentially resulting in negotiation of meaning), but the frequency of suprasegmental issues relative to other pronunciation features was not considered. Nevertheless, those suprasegmental features were also identified as functionally useful for discourse management in negotiation for meaning that followed misunderstandings. In sum, these studies, which examined interaction either outside of the classroom or from a noninteractionist perspective, underscore the potential role that pronunciation can play during task-based interaction.

Focus on Pronunciation during Classroom Language-Related Episodes

By investigating LREs, which involve learner topicalization of linguistic forms during interaction (Swain & Lapkin, 1998), researchers have documented the extent to which specific areas of language are targeted during task-based interaction. Typically, learners focus most

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

on vocabulary during peer interaction (e.g., Bowles, Toth, & Adams, 2014; Williams, 1999); however, several studies have reported some attention to pronunciation as well. These studies tend to code the linguistic focus of LREs in broad terms, grouping any phonological trigger, segmental or suprasegmental, in the same category (e.g., Bowles, Toth, & Adams, 2014; Zhao & Bitchener, 2007). However, some studies have not specifically stated what comprised phonological LREs (e.g., Bueno-Alastuey, 2013), while at least one (Kennedy et al., 2015) coded segmental and suprasegmental LREs separately. Due to these differing operational definitions of pronunciation-related LREs, caution is needed when considering cross-study comparisons.

Nevertheless, on the higher end, Bueno-Alastuey (2013) found an overall 40% rate of pronunciation-related LREs, while Zhao and Bitchener (2007) reported that pronunciation was the focus in 28% of all LREs. Bueno-Alastuey, in the only CMC study considered here, compared NNS dyads in which interlocutors either shared an L1 or had different L1s, along with NNS-NS dyads (14 dyads for each type), in audio-only SCMC. She found that pronunciation received the largest share of attention in NNS-NS pairings (49%) and the least in different-L1 pairs (35%), but it should be noted that, in terms of raw numbers, those different-L1 pairs generated more pronunciation LREs (54) and LREs in total (155 compared to 85 in the NNS-NS dyads). Zhao and Bitchener's study, which featured 16 diverse adult ESL learners in 10 hours of task-based classroom interaction, revealed that learner-learner interactions focused on pronunciation proportionally more than did teacher-learner interactions (28% vs. 17%, respectively). In another study, Kennedy et al. (2015) identified pronunciation as the focus of 18% of LREs in interactions between intermediate and advanced French learners in Canada. Digging deeper into the source of pronunciation-related LREs than most other studies, Kennedy et al. (2015) also categorized the pronunciation LREs as being segmental (stemming from vowel or consonant errors) or suprasegmental (such as liaison or stress errors). It was found that most intelligibility issues were related to segmentals (n = 21), with suprasegmentals (n = 3) rarely leading to breakdowns in understanding. In sum, these studies explored learner-learner interactions in which pronunciation played a relatively prominent role in communication and negotiation for meaning.

Other studies have found pronunciation LREs to be considerably less frequent. For example, Gurzynski-Weiss and Baralt (2014) observed a focus on pronunciation in only 3% of LREs in FTF Spanish as a foreign language task-based interaction. Similarly, in their study of Spanish L2-L2 and L2-Heritage Learner (HL) pairings, Bowles et al. (2014) found pronunciation LRE rates of 1% for L2-L2 and 4% for L2-HL dyads. However, learners self-corrected their pronunciation at noticeably higher rates: 8% for L2-L2 pairs and more than 19% for L2-HL pairs, suggesting

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

that pronunciation was not entirely ignored during these interactions. These conflicting reports on the frequency of pronunciation-related LRE call for further investigation and suggest that moderating variables could be at play.

L1 and Task: Moderating Influences on Pronunciation

L1 background and task demands are two factors known to affect L2 pronunciation. L1 background is linked to L2 speech perception and production in prominent frameworks of L2 phonological acquisition (e.g., Eckman's [2008] Marked Differential Hypothesis; Flege's [1995] Speech Learning Model), providing a theoretical basis for the notion that L2 pronunciation and speech perception can be moderated by L1 background. Empirical evidence has demonstrated an effect of L1 on speech comprehensibility (Crowther et al., 2015), perception of speech sounds (Ortega-Llebaria & Colantoni, 2014), and judgments of comprehensibility (Saito & Shintani, 2015). Interestingly, NS listeners who have knowledge of an L2 speaker's L1 are often more sympathetic in their judgments of L2 speech (Winke, Gass, & Myford, 2013), and in an early study, Gass and Varonis (1984) found that NS familiarity with L2 speech in general, particular L2 accents, and individual L2 speakers facilitated comprehension. Evidence accumulated so far about homogenous and heterogeneous learner pairings in interaction, however, is ambiguous: Bueno-Alastuey (2010, 2011, 2013) found significantly different amounts of pronunciation LREs between same-L1 and different-L1 dyads, whereas Kennedy et al. (2015) found no difference in pronunciation-related LREs across the two pairings. The former finding is corroborated by several aforementioned speech studies, which suggest L1 familiarity matters. The latter finding draws support from Munro, Derwing, and Morton (2006), where L2 speakers from diverse backgrounds generally agreed on the intelligibility, comprehensibility, and accentedness of diverse L2 speakers (with very small between-groups differences) and did not consistently find the intelligibility of speakers from shared L1 backgrounds to be higher.

Research has shown that task conditions can also influence interaction (e.g., Ellis, 2003; Keck et al., 2006); however, to our knowledge, no studies have compared the effects of different task conditions specifically on pronunciation. Consequently, a brief review of lexical and morphosyntactic research is considered. One important task condition is the direction of information flow, with tasks being either one way, in which one learner possesses all the task information, or two way, in which learners each have different pieces of information (Pica, Kanagy, & Falodun, 1993). There is consensus that two-way tasks produce considerable negotiation in both FTF (Ellis, 2003) and CMC (e.g., Blake, 2000) interaction.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Another important task condition is the linguistic focus (Ellis, 2003; Smith, 2004). Tasks may be either unfocused, with no specific linguistic targets, or focused, with learners expected to use certain linguistic structures. With focused tasks, researchers and teachers try to ensure that certain linguistic structures are task essential by trying to necessitate learners using the language forms to complete the task. Additionally, tasks may be seeded with the targeted linguistic forms to encourage their use. For example, Smith (2004) seeded his tasks with low-frequency lexical items to encourage learner vocabulary-related negotiation. Task essentialness was one of the task conditions that Keck et al. (2006) found to influence L2 development.

Finally, the expected task outcome is important for negotiation (Ellis, 2003; Smith, 2004). In closed tasks, there is only one correct outcome, while open tasks can have multiple correct outcomes. For example, in picture narration tasks, there is generally only one correct order of events; however, in decision-making tasks, learners are often able to negotiate the outcome, such as which items are most important to take on vacation. Another component of task outcome is whether tasks are convergent, with learners needing to agree to successfully complete the task, or divergent, in which learners do not need to agree. In a decision-making task, for example, learners may be required to converge and agree on the most important items, or they may be allowed to diverge and simply discuss which items are most important in their own opinions. In general, tasks that necessitate agreement on a single correct outcome generate more interaction (e.g., Ellis, 2003; Gass, Mackey, & Ross-Feldman, 2005).

Although task-based research has not compared the effects of different task conditions on pronunciation, several studies have used focused tasks to investigate specific phonological features. For example, Sicola (2008) analyzed the interaction of adult ESL learners in a map-related task designed to elicit /θ/. For 26 dyads, she observed 559 corrective feedback moves that occurred during negotiation of meaning related to location names. This resulted in 186 instances of learner modified output (roughly seven per dyad), of which learners achieved more targetlike pronunciation 65.1% of the time. Although the importance of English interdental fricatives for listener understanding is generally considered to be quite low (Munro & Derwing, 2006), Sicola's results nonetheless suggest that focus on phonological forms in interaction can be beneficial for learners. Looking at a different English segmental feature with greater importance to intelligibility, Saito (2015) presented groups of Japanese learners of English with tasks that required accurate perception and production of words with an initial /ɹ/ to investigate the effects of form-focused instruction and corrective feedback. Satio found that learners improved in their perception and production of the target phoneme. Together, these studies suggest that learner-learner interaction can be an effective context for a focus on pronunciation.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

COMMUNICATION MODALITY AND PRONUNCIATION

With the growth of Internet communication technologies, mode of communication has represented a growing area of interest in L2 research (Sauro, 2011; Smith, 2004). The modes of CMC made available by technology have direct impacts on instructional affordances in several contexts, including (but not limited to) distance learning, teletandems, language learning social networks (e.g., *Busuu*, www.busuu.com), and hybrid or "flipped" courses (Blake, 2011).

Modality and its effects on L2 use and acquisition have long been a concern, as the earliest forms of CMC were restricted to asynchronous text-based communication that is fundamentally different from the in-person communication that has driven interactionist SLA theory and characterizes many ISLA contexts. FTF interaction is inherently multimodal, with speakers and listeners taking full advantage of gestures, visual aids, and the immediate environment (Chamberlin Quinlisk, 2008). Although the FTF modality, with audio and visual channels, is considered the baseline for modality comparisons, it should not necessarily be assumed that having two (or more) channels is beneficial for learning: Each channel to be processed can increase the cognitive load for learners and potentially lead to overload (Mayer & Moreno, 2003; Yanguas, 2010). Nowadays, SCMC also presents an array of modal possibilities, potentially combining video, text, and audio. In terms of general learning outcomes, Ziegler's (2016a) meta-analysis of 14 studies comparing FTF and SCMC interaction found a small effect size (d = .13) in favor of SCMC; however, the 95% confidence intervals crossed zero, indicating that there is no statistical superiority of one condition over the other. Lin's (2015) meta-analysis of oral proficiency outcomes in 25 CMC and FTF interaction studies obtained a somewhat larger effect size of .40. These findings may be encouraging for the use of SCMC in ISLA, although they do not speak directly to the qualities of SCMC interactions nor are they clear on whether the different types of SCMC (text only, audio only, or audiovisual) and FTF modalities differentially affect focus on pronunciation.

Studies investigating various modes of SCMC have suggested that audiovisual CMC contains interaction similar to FTF interaction, whereas text-only CMC tends to be quite different (Blake, 2000; Jepson, 2005; Sauro & Smith, 2010; Yanguas, 2010) and naturally excludes pronunciation as a factor. For example, modality can influence the quantity and types of LREs. Jepson (2005) found more repair moves in audio-only CMC than written, with many of the audio-only repairs focused on pronunciation. Yanguas (2010), comparing FTF, audiovisual CMC, and audio-only CMC task-based interactions, found that the quantity and quality of LREs in FTF and audiovisual CMC were similar, while audio-only CMC yielded a larger quantity of LREs that involved more linguistic elaboration.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Yanguas argued that when audio is the only mode available for interaction, contextual information potentially useful for interpreting meaning is lost. Therefore, Yanguas suggested that learners in the audio-only condition were less successful in establishing understanding when negotiating for meaning because of the loss of the visual channel and inability to use visual cues. Without visual support, learners may need to rely more on linguistic features, such as phonological form, to maintain or repair understanding.

Particularly germane to pronunciation and modality is the impact of visual information on phonological perceptions. Examining the McGurk effect (the influence of lip-movement visuals on phonemic perceptions), Hardison (1999) found that L2 English learners who received visual information that matched audio were more successful in phoneme recognition than those who were presented with conflicting visuals (e.g., visuals of /fa/ presented with the sound /pa/). In another study, Sueyoshi and Hardison (2005) reported that both low-intermediate and advanced learners of English benefited in terms of comprehension when audio was presented with video showing the face of speakers, with the low-intermediate learners benefiting further from video showing the speakers' faces and gestures. In sum, having access to visual information is likely to benefit speech intelligibility and presumably reduce the likelihood of phonologically-related communication breakdowns, a prediction that finds some initial support from the relatively high frequencies of pronunciation-related LREs in the audio-only conditions of Bueno-Alastuey (2010, 2013) and Jepson (2005). Because of the ever-increasing pedagogical use of technology for communication both inside and outside the classroom, it is important for L2 teachers and ISLA researchers alike to better understand the effects of modality on interaction, especially in regard to LREs, including pronunciation-related ones, that potentially affect L2 development.

RESEARCH QUESTIONS

Motivated by the relative lack of focus on pronunciation in investigations of both FTF and SCMC learner-learner interactions and a need to account for the potentially moderating influences of L1 and task, the present study is guided by the following research questions (RQs):

RQ1. How frequent are pronunciation LREs in L2 learner task-based interaction?
  RQ1.1. How frequently do segmental and suprasegmental pronunciation features trigger pronunciation-related LREs in L2 learner task-based interaction?
RQ2. How does L1 pairing (same-L1 vs. different-L1) influence the frequency of pronunciation LREs during L2 learner task-based interaction?

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

RQ3. How does modality (FTF vs. audio-only SCMC) influence the frequency of pronunciation LREs during L2 learner task-based interaction?
RQ4. How do task features influence the frequency of pronunciation LREs during L2 learner interaction?

In support of the research questions, the following hypotheses were formed.

RQ1. Based on previous research finding minimal focus on pronunciation during L2 interaction, it was hypothesized that pronunciation LREs would be limited throughout the interaction.
  RQ1.1. The limited research in this area suggests that most pronunciation LREs would be triggered by segmental, rather than suprasegmental, features.
RQ2. Because learners are more familiar with their own L1 accents, it was hypothesized that same-L1 dyads would engage in fewer pronunciation LREs. However, the effect of L1 familiarity might be limited due to learners' exposure to a variety of L1 accents in the second language context in which the study was conducted.
RQ3. Due to research that indicates that nonverbal cues can play a role in phonological processing, it was hypothesized that more pronunciation LREs would occur in audio-only SCMC.
RQ4. Previous task-based research suggests that tasks with (a) more task essential language, (b) convergent outcomes, and (c) closed outcomes will contain more LREs.

METHODS

Methodological Framework

The current study, based on a subset of data from Loewen and Wolff (2016), involves the observation and description of learner-learner interaction during communicative tasks. As such, no intervention was provided to direct learners' attention to specific linguistic structures; rather, learners were instructed to focus on completing tasks, which are common to L2 classroom contexts. A between-groups design was employed for modality, and a within-groups design was used for task type. A convenience sample was used for L1 pairing.

Participants

The participants in this study (N = 30) ranged in age from 17 to 35 with an average age of 22 (SD = 5.5), and had begun English study at an

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

average age of 13 (SD = 7.8). The participants were enrolled in intermediate-level ESL classes at a large midwestern university in the United States. Participants reported using English 8 hours (SD = 7.7) per week outside of class, and their self-rated proficiency on a 5-point scale for the four L2 skills was around 3: speaking (M = 3.5, SD = 1.1, range 1–5), listening (M = 3.3, SD = .96, range 1–5), reading (M = 3.0, SD = .80, range 1–4), and writing (M = 2.9, SD = .80, range 1–4). For logistical reasons, no additional proficiency testing was conducted nor were standardized test scores, such as TOEFL or IELTS uniformly available; therefore, the researchers had to rely on institutional placement as a proxy for L2 proficiency despite its inexact qualities (Thomas, 1994). Participants' L1s were Chinese (n = 17), Arabic (n = 10), Korean (n = 2), and Cantonese (n = 1). Participants were asked about their frequency of use of "oral voice-chat," with 1 indicating "Never" and 5 "Very Often." A mean of 3.8 (SD = .72), with a range of 2 to 5, indicates that all participants had at least minimal exposure to oral SCMC modalities.

Materials

The participants completed three different tasks: a picture differences task, a consensus task, and a conversation task. The tasks, which are presented in the following text, were chosen because they are representative of interactive tasks used in L2 classrooms (e.g., Basturkmen, Loewen, & Ellis, 2004) as well as L2 research (e.g., Gass, Mackey, & Ross-Feldman, 2005; Mackey, 1999). Each dyad was randomly assigned to one modality and completed all three tasks in that modality; thus, modality was a between-groups variable, whereas task was a within-groups variable. All three tasks allowed learners to use their own linguistic resources, were related to real-world activities, and had nonlinguistic outcomes (Ellis, 2003).

Picture Differences Task. In this task, participants had to spot the difference between two similar pictures. The pictures were the same ones used by Mackey (1999) and Gass et al. (2005, pp. 608–609) and depicted nearly identical park scenes with several differences between the pictures. For example, in one version, a girl is playing with a ball, while in the altered version the ball is missing. Participants had to share information about their respective pictures to complete the task. No linguistic information was provided, and learners had to agree on a specified set of differences.

Consensus Task. This task, in Appendix A, required participants to come to an agreement about the information provided to them, based on a set of criteria. The task was conceptually similar to consensus tasks observed by Basturkmen et al. (2004) in ESL classrooms and used

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

in Gass et al.'s (2005) research; however, the content differed. Instead of deciding which prisoner should be released early on parole or where a hypothetical friend should attend college, the dyads had to exchange information about three scholarship candidates and decide which candidate should receive a scholarship to study in the United States. Participants were given a sheet with information about the scholarship requirements, which they were told to consider when deciding on the scholarship recipient. Participants also received different pieces of information about each candidate, which they needed to exchange to come to a consensus on one candidate. The information was designed so that all candidates had positive and negative qualities; thus, learners could potentially choose any candidate for the scholarship.

Conversation Task. To elicit more naturalistic interaction between participants (Nakahama, Tyler, & van Lier, 2001), a conversation task, in Appendix B, was used in which participants asked about their partners' favorite things, in addition to describing their own. Each participant received a list of discussion topics, which contained the same topics but in different orders. Participants were instructed to take turns asking each other about their favorites (e.g., restaurant, film, book) and to elaborate whenever possible. Furthermore, learners were asked to discuss commonalities and differences between their answers.

All three tasks were two-way information gap tasks to maximize negotiation; however, several other task conditions were varied, as shown in Table 1. The Scholarship task was the only task in which the researchers provided linguistic information to be exchanged. In the two other tasks, participants relied on their own linguistic resources. The Favorites task was the only divergent task in which learners did not have to agree with each other to complete the task. Finally, the Differences task was the only closed task, with only one correct outcome.

Procedure

Participants were recruited from intact classes and invited to participate in an English interaction study outside of class time. Learners were

Table 1. Task characteristics

Task	Linguistic Input	Convergent	Open
Differences	-	+	-
Favorites	-	-	+
Scholarship	+	+	+

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

not told in advance that there would be different interaction modalities, and when they came to the lab, learners were randomly assigned to one of the modalities, resulting in eight FTF and seven audio-only CMC dyads. The order of the tasks was counterbalanced. Before beginning each task, one of the researchers explained the task to the two participants, and after a maximum of 20 min on a task, the researcher stopped the participants (if they had not finished earlier) and moved on to the next task. Each session lasted a total of 50 to 60 min. For the FTF modality, a camera was used to record the participants who were seated across from each other. In the audio-only CMC modality, participants were in separate rooms seated in front of a computer. Skype was used for the audio-only CMC groups, but for logistical reasons, only audio was available; consequently, the participants did not see each other, an occurrence not uncommon in SCMC (Bueno-Alastuey, 2013). Audio recordings of the CMC interaction were saved for subsequent analysis. At the end of the three tasks, the participants filled out a background questionnaire and were paid 20 U.S. dollars each for their participation.

Learners self-selected their partners for the interaction; therefore, we were not able to control for L1 background. However, the self-selected pairings included both same-L1 (n = 11) and different-L1 (n = 4) dyads (Table 2). On the background questionnaire, participants indicated their level of familiarity with their conversational partner, with 1 specified as "'Don't know at all" to 5 "I know him/her very well." The mean score was 3.6 (SD = 1.1), with a range from 2 to 5, indicating at least minimal but more often considerable interlocutor familiarity.

Coding

To identify attention to language items, interactions were coded for LREs, which are "any part of a dialogue in which students talk about language they are producing [or] question their language use [or] other- or self-correct" (Swain & Lapkin, 1998, p. 70); however, unlike Swain and

Table 2. L1 Pairings by languages and mode

L1s	SCMC	FTF
Arabic-Arabic	3	1
Chinese-Chinese	2	5
Arabic-Chinese	1	1
Chinese-Korean	1	0
Cantonese-Korean	0	1
Total	7	8

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Lapkin, we did not include self-corrections in our analysis because our primary interest was the interaction that occurred between participants. While potentially providing evidence that an erroneous linguistic item has been noticed, self-corrections do not necessarily occur because of learner-learner interaction. Another characteristic of LREs is that they are often quite explicit in their focus on language; however, we also included more implicit types of attention to language, such as negotiation for meaning and recasts (e.g., Long, 1996). Occurrences of interaction moves (confirmation checks, clarification requests, comprehension checks, LREs, and recasts) were transcribed and coded by the first author and a research assistant (see Appendix C for transcription key).

Once a pronunciation LRE was identified, it was then secondarily coded for type of phonological feature, either segmental or suprasegmental. Individual phonemes (vowels, consonants) and multiple phonemes were coded as segmental. Suprasegmental LREs included prosodic features, such as intonation, word stress, and sentence-level stress. An "other" category was used on three occasions for pronunciation LREs clearly involving both segmental and suprasegmental features or that otherwise could not be reconciled.

One way in which a pronunciation focus could happen was through overt topicalization by one or both learners. In Example 1, learner B requests clarification of learner A's assessment of the number of people in her picture. The clarification request indicates that A's pronunciation of *thirteen* was unintelligible, due to poor production by A and/or poor perception by B because B did not produce the correct number but rather a phonologically close approximation, *thirty*. This LRE was coded as segmental because the final vowel and consonant appear to trigger comprehension difficulty. Interestingly, A uses a suprasegmental strategy, increasing her volume and stressing the final syllable of *thirteen*, to respond to the segmental clarification request.

Example 1. Differences task (audio only)

(counting people in picture)
A: thirteen just
B: thirty?
A: thirTEEN thirTEEN (closer to mic)
B: one three?
A: one three yeah (.) thirteen
B: °okay°

Another example of a pronunciation LRE is when negotiation for meaning occurs due to nontargetlike pronunciation. In Example 2, learner B mispronounces the segment /f/ in *fence*, leading to a breakdown in communication that is resolved only after she spells the problematic word.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Example 2. Differences task (audio only)

A: Do you have anything in the left bottom?
B: Left?
A: Left corner?
B: oh I have uh some [p^{h}æns]
A: What is [p^{h}æns]?
B: Like a protector (.) the tree
(..)
B: I think it's [p^{h}æns]
A: How to spell?
B: F-E-N-C-E
A: F (.) E-N-C oh fence (.) okay
B: yeah

The final type of pronunciation LRE involved other-correction, either implicit or explicit, of a nontargetlike pronunciation even if no obvious breakdown in communication occurred. In Example 3, learner A mispronounces the initial consonant /p/ and vowel /ɛ/ in *pet* (a segmental error), and in response, learner B produces the correct form.

Example 3. Favorites task (audio only)

A: so do have (.) do you like uh bit (.) do you have bit?
B: I don't have [pet but
A: [ca:t cat
B: mm I don't have because in my dormitory they didn't allow us to=
A: =ooh
B: to have a pet

At times, it was difficult to determine if pronunciation issues were implicated in the interaction. Thus, if a clarification request or other type of negotiation for meaning did not involve a difference in pronunciation between the two learners, no pronunciation LRE was counted, even though the pronunciation may not have been completely targetlike. In Example 4, learner B produces a confirmation check in response to learner A; however, there is no discernible difference in pronunciation. Therefore, although this interaction constitutes an LRE, it was not counted in the current analysis because it did not topicalize phonological features.

Example 4. Differences task (FTF)

B: how many (.) girls you have with the pink dress?
A: uh just one
B: just one?
A: yeah

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

To establish the reliability of pronunciation LRE coding, the entire dataset was double coded. Where there was a disagreement in coding decisions, the researchers engaged in a discussion until a consensus was reached. For the segmental/suprasegmental coding, 25% of the data were double coded with an interrater agreement of 100%.

Analysis

To answer the research questions, the number of pronunciation LREs was tallied for each dyad in each task. Due to the generally low frequency of pronunciation LREs, no inferential statistics were calculated. Furthermore, the data are not independent, and therefore do not meet one assumption of statistical tests, such as chi-square. However, 95% confidence intervals are presented to indicate the level of measurement precision, which provides more information and is preferable to null hypothesis significance testing (e.g., Norris, 2015; Plonsky, 2015). When comparing two sample means and their confidence intervals, a difference in population means can be inferred if the confidence intervals have no overlap. Overlaps in confidence intervals do not exclude the possibility of a difference in the population but do introduce the possibility of there being no difference at all, and thus differences between means with overlapping confidence intervals ought to be interpreted conservatively and cautiously.

Results

To answer the first question regarding the frequency of pronunciation LREs during the interaction, it is first necessary to determine the overall number of LREs. There was a total of 1,114 LREs across all observed interactions, with an average of 74.2 LREs per dyad (SD = 35.02) and a range from 22 to 134.

Having established the total number of LREs, it is possible to determine the number of pronunciation-related LREs during the task-based interactions. Overall, 158 of the 1,114 LREs targeted pronunciation, with an average of 10.5 (95% CI [6.3, 14.7], SD = 7.5) per dyad and a range from 2 to 29. In terms of percentage of total LREs, there was an average of 14.1% pronunciation-related LREs (95% CI [10.4, 17.9]) per dyad, with a range of 5% to 24%. The remaining LREs focused on either grammar or vocabulary.

Related to the first research question, these pronunciation LREs were also coded broadly for the type of phonological feature focused on.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Out of 158 pronunciation LREs, 143 (90%) focused on segmental features and 12 (8%) on suprasegmental, with 3 LREs classified as "other."

The next research question asks about the effects of L1 background on pronunciation. Table 3 shows that 12% of LREs targeted pronunciation in same-L1 dyads, whereas 19% occurred in different-L1 dyads. The 95% confidence intervals overlap, suggesting that there may not be a difference in the population mean. However, on average, the L1-different pairs produced nearly twice as many pronunciation LREs as the L1-same pairs.

The next research question addressed the role of modality. Table 4 indicates that 11% of LREs in audio-only SCMC targeted pronunciation, whereas 16% of FTF LREs were pronunciation related. The 95% confidence intervals overlap considerably, suggesting that there may not be a difference in the population mean.

The final research question pertains to the distribution of pronunciation LREs during the different tasks. Table 5 reveals that the Differences and Favorites tasks had very similar levels of pronunciation focus at roughly 10%. However, the Scholarship task had more than twice as much pronunciation focus, at about 20%. Nevertheless, because of the overlapping confidence intervals, it is not possible to state that the population mean for the Scholarship task is significantly different from the other tasks.

DISCUSSION

The first research question asked about the amount of focus on pronunciation during learner-learner interaction. The results showed that, on average, roughly 16% of LREs during interaction were pronunciation focused. These results are lower than Bueno-Alastuey's (2013) 40%, and Zhao and Bitchener's (2007) 28%, but higher than the 1 to 4% found in Gurzynski-Weiss and Baralt (2014) and Bowles et al. (2014), respectively. The levels of attention to pronunciation in the current study are most similar to Kennedy et al.'s (2015) rate of 18%.

It is not immediately clear how to interpret the current data in relation to the diverse percentages found across other studies. Different criteria for identifying pronunciation LREs may account for some of the

Table 3. Pronunciation LREs in L1-same and L1-different dyads

		95% CI				95% CI		
Dyad	*M*	Lower	Upper	*SD*	%	Lower	Upper	*SD*
L1-Same	8.5	3.2	13.8	7.8	12.3	8.0	16.4	6.2
L1-Different	16.0	13.4	18.6	1.6	19.4	10.6	28.2	5.5

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Table 4. Pronunciation LREs across modes

		95% CI				95% CI		
Mode	*M*	Lower	Upper	*SD*	%	Lower	Upper	*SD*
Audio-only CMC	7.5	1.9	12.9	6.0	11.5	5.0	18.0	7.1
FTF	13.2	6.6	19.9	7.9	16.5	11.6	21.4	5.9

variability across studies.[1] Some studies explicitly included segmental and suprasegmental features in their coding criteria (Kennedy et al., 2015; Zhao & Bitchener, 2007) as in the present study, but others were less specific (Bowles et al., 2014) or more restrictive (Bueno-Alastuey, 2013) in their criteria.

In addition to methodological factors, it may be the case that different learner characteristics and interaction contexts are responsible, in part, for the varied results across studies. For example, the participants in the current study shared some similar characteristics with those in Kennedy et al. (2015), though the target languages were different. Both studies involved young adult learners enrolled in intermediate-level L2 course in a second language context engaging in information gap tasks. In contrast, Bueno-Alastuey's learners were foreign language learners who were based in their home countries and who communicated using Skype with L2 speakers in other countries, while the learners in Gurzynski-Weiss and Baralt (2014) and Bowles et al. (2014) were L2 learners of Spanish studying at American universities. It could be that learners' presence in a larger L2 speaking context might heighten their awareness of pronunciation issues (Derwing & Munro, 2015), whereas learners in foreign language contexts, surrounded by compatriots with similar accents, may feel less need to focus on pronunciation during interaction. Further research is needed to explore the relationship between second language versus foreign language contexts and L2 learners' overall awareness of linguistic issues, in general, and pronunciation issues, in particular. However, most studies, including the current one, suggest that during learner-learner interaction, pronunciation is topicalized considerably less than other aspects of language such as grammar and vocabulary.

Table 5. Pronunciation LREs across tasks

		95% CI				95% CI		
Task	*M*	Lower	Upper	*SD*	%	Lower	Upper	*SD*
Differences	2.67	1.54	3.79	2.02	12.8	7.1	18.6	10.4
Favorites	2.73	.90	4.56	3.30	9.7	5.4	14.0	7.7
Scholarship	5.13	2.42	7.84	4.90	20.6	11.6	29.7	16.3

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

In terms of the two general categories of phonological features that triggered pronunciation-related LREs, segmental features were by far the most prevalent at 90%, suggesting that intelligibility during L2 interaction was disproportionately affected by the mispronunciation of individual consonants and vowels. Again, these results are in line with Kennedy et al. (2015), which also found an overwhelming majority of pronunciation LREs focusing on segmentals in L2 French interaction. The current results also align with Matsumoto's (2011) descriptions of NNS-NNS conversations. Considering L2 pronunciation more broadly, the apparent importance of segmental features aligns with Jenkins's (2002) description of a core set of phonological features crucial for successful English as a lingua franca communication. At the same time, the present findings could be interpreted to challenge the emphasis some researchers place on suprasegmental features as pedagogical targets (e.g., Derwing, Munro, & Wiebe, 1998). However, Levis (2016) argues that the research supporting the superiority of suprasegmentals is scarce; furthermore, as Jenkins argues for segmentals, not all suprasegmental features may be equally influential in hindering intelligibility. Therefore, we interpret the current findings as being reflective of the importance of segmentals in this specific communicative context: The importance of suprasegmental features in communication featuring extended discourse and longer turns, such as lectures or narration, may be diminished in the shorter turns and rapid exchanges of classroom task-based interaction or social conversation. Nonetheless, it would appear that segmental features are worthy of pedagogic attention in the context of interactive classroom tasks, and furthermore it may be helpful to teach strategies for working through segmental-based lapses in intelligibility (such as those employed by the skilled communicators in Matsumoto [2011]).

Regarding the factors moderating the frequency of pronunciation-related LREs, there were no clear differences within any of the three variables: L1 background, task modality, or task conditions. For L1 background, it is not possible to state with confidence that there were meaningful differences between same-L1 and different-L1 dyads in the current study, even though the overall number of pronunciation LREs was twice as high among the different-L1 pairs. Caution is needed in interpreting these differences due to the small number of different-L1 dyads ($n = 4$), and the slightly overlapping 95% confidence intervals. Still, the current results lend some support to the notion that familiarity with one's own L2 accent may decrease the need for negotiation for meaning that is triggered by pronunciation issues, as evidenced in Bueno-Alastuey's (2013) finding of statistically more pronunciation-related negotiation in different-L1 dyads and Gass and Varonis's (1984) conclusion that familiarity with an L2 accent facilitates comprehension. The ambiguous result in the current study may be affected by the heterogeneity of L1 background in the ESL classrooms, such that even if learners were speaking

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

with same-L1 partners, learners may have become accustomed to other L1 accents (Munro et al., 2006). However, with the small and unequal number of same-L1 and different-L1 dyads, this issue warrants further research.

Regarding modality, there were no differences in the number of pronunciation-related LREs between the FTF and audio-only CMC dyads, even though it was hypothesized that the lack of visual cues in the audio-only CMC would increase the frequency of negotiation for meaning based on phonological issues (Sueyoshi & Hardison, 2005; Yanguas, 2010). Although Jepson (2005) found more linguistic elaboration in negotiation of meaning during audio-only CMC compared to text SCMC and that a considerable portion of audio-only LREs were related to pronunciation, there was no comparison to FTF interaction, making the findings difficult to relate to the present study and inconclusive regarding the role of the visual channel. Yanguas (2010), in a study that compared FTF and audio-only CMC and included a consensus task, reported frequent use of visual cues in the FTF mode, while audio-only CMC generated more linguistic elaboration. However, Yanguas's study focused on lexical negotiation for meaning, and it may be that phonological breakdowns in communication do not rely on visual cues, such as pointing or gesturing, to the same degree. Indeed, research that has found positive effects of visual cues on pronunciation perception (e.g., Sueyoshi & Hardison, 2005) investigated monologic production, rather than dialogic interaction. Another interpretation could be that the loss of the visual channel is counterbalanced to some degree by the lower cognitive load of processing information from a single channel (Mayer & Moreno, 2003). Continued investigation into the role of the visual channel as it relates to pronunciation in interaction would be informative.

Regarding task conditions, there were noticeable, but not statistical, differences in the number of pronunciation-related LREs among the three tasks, with the consensus task (i.e., Scholarship) averaging roughly twice as many pronunciation LREs as the other two tasks. Although this difference did not pass the 95% confidence interval threshold, it is interesting to consider the possible reasons for the visible differences. First and foremost, the notion of task-essential language (Keck et al., 2006) comes to mind. Although there were no specific target structures in the Scholarship task, it was nevertheless "seeded" with linguistic items, in terms of the candidates' information, that needed to be exchanged to successfully complete the task. In contrast, the pictures in the Differences task provided only nonlinguistic information, and the prompts in the Favorites task consisted of one or two words that served as topics for discussion rather than information that needed to be exchanged. Recent examples of tasks seeded with exemplars of phonological targets include Saito's (2015) interventionist investigation, which used tasks seeded with English /ɹ/ to improve Japanese learners' perception and

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

production of the phoneme. Learners were provided with discussion prompts such as "Running inside is better than running outside" and "Is reading comic books good for children?" (p. 385). Similarly, Sicola's pedagogical intervention (2008) seeded minimal pairs involving /θ/ (e.g., *Boot Road* and *Booth Road*) into a closed-outcome map task (p. 143). Learners were required to ask each other for directions to specific locations, and the geographic proximity of intersections with minimal pair names demanded phonological accuracy for successful task completion.

In a post hoc analysis to explore this possibility, it appeared that many of the pronunciation LREs in the Scholarship task occurred around the specific information provided for the learners. In Example 5, learner B is trying to convey that one candidate had *A grades* in every subject; however, instead of stressing *A* to indicate a score, the learner produces a schwa vowel as is usual for an indefinite article. Learner B also mispronounces *grades*. Both these issues led to negotiation around pronunciation and meaning. If the learners had not been provided with the phrase *A grades in every subject*, we can speculate that the learners would have been less likely to negotiate the candidate's scholastic ability. Of course, further evidence to support this hypothesis is needed; however, this observation does indicate that if teachers wish for learners to focus on comprehensible pronunciation, they may want to provide tasks that require the production of specific linguistic forms rather than leaving learners entirely to their own devices, whereby they might be able to avoid difficult words and phrases.

Example 5. Scholarship task

B: and the the second is a grands in every subjects
A: A?
B: A grands in every subjects means (.) he grades A
A: a grade A is very good
B: yeah
A: yeah I understood

Limitations

Even though the current study provides insight into the limited focus on pronunciation, as well as the limited influence of moderating variables, during learner-learner task-based interaction, these findings must be considered in light of the study's limitations. First, the small sample size (30 learners in 15 dyads), particularly in the L1 pairings (only two L1-different dyads per modality), reflects a well-known limitation in SLA research. Second, while the participants were recruited from intermediate ESL classes, their length of English study varied considerably, and no additional proficiency measures were obtained. Third, the small

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

sample size and dependent nature of the data precluded performing inferential statistical tests to locate and measure any generalizable effects. Fourth, the research design was opportunistic in nature, and as such lacked a quasi-experimental control. For example, no intervention was used to assign learners to specific dyads. Instead, learners self-selected their partners and, as is typically the case in L2 classes, they may have been able to choose someone familiar to them or they may have had to work with someone they did not know very well. In the current study, it is more likely that they chose familiar partners because they had to participate in the research outside of class time; however, because no information was collected pertaining to learners' decisions to participate in the study, it is not possible to know for certain. Similarly, it is likely that learners who were more motivated, whether by money or a desire to practice their English, participated in the study. Nevertheless, the participants were selected from existing L2 classes, and the method of dyad grouping reflects a common pedagogic practice, thereby maintaining ecological validity, although at some expense to experimental control. Fifth, no introspective measures were utilized to collect learner perspectives on L1 backgrounds, modality, task characteristics, or their own orientation to L2 pronunciation. Analyzing learner views on their own intelligibility and that of their interlocutor would have been informative. Finally, to reduce the amount of time each dyad spent interacting, it was decided to use a between-groups design for modality and a within-groups design for task. Thus, it is possible that individual differences may have had a larger impact on the interaction in the different modalities, but a more minimal impact due to task type. However, the background questionnaire inquired about participants' frequency of online SCMC (both audiovisual and written); results showed that learners were generally familiar with SCMC, suggesting that any technological novelty did not unduly influence the interaction in that modality. Additionally, no learners expressed concern about interacting in the audio-only SCMC condition.

CONCLUSION

Despite these limitations, the current study provides useful evidence of the state of affairs during learner-learner interaction based on communicative tasks that are commonly used in task-based research and instruction. The analysis of pronunciation in this context adds to the limited base of empirical knowledge regarding this linguistic area, especially compared to the greater amount of research on grammar and vocabulary in task-based research. However, because of the variability of findings in previous studies, more descriptive research is necessary to gain a clearer picture of pronunciation and task-based interaction.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Additionally, interventionist quasi-experimental research is needed to determine the effectiveness, if any, of such interaction on L2 pronunciation development.

Finally, for practitioners, these findings support the use of task-based interaction as a means of addressing phonology in both FTF and oral CMC contexts. Additionally, the results suggest that teachers would do well to focus on segmental features, which can interfere with intelligibility, particularly in interactive communication. To increase the likelihood of a phonological focus, teachers may wish to target important segmental features in the design of interactive tasks. In this way, pronunciation, which is important for L2 intelligibility, can be incorporated into activities that may already be in use in the classroom. As a result of a focus on pronunciation during task-based interaction, L2 learners may be able to improve their ability to be understood, a goal not inconsequential for L2 learning.

Received 26 February 2016
Accepted 5 November 2016
Final Version Received 8 November 2016

NOTE

1. We are grateful to an anonymous reviewer for suggesting this possibility.

REFERENCES

Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 30–51). Oxford, UK: Oxford University Press.

Basturkmen, H., Loewen, S., & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics*, *25*, 243–272.

Blake, R. J. (2000). Computer mediated communication: A window on Spanish L2 interlanguage. *Language Learning & Technology*, *4*, 120–136.

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

Bowles, M. A., Toth, P. D., & Adams, R. J. (2014). A comparison of L2–L2 and L2-heritage learner interactions in Spanish language classrooms. *Modern Language Journal*, *92*, 497–517.

Bueno-Alastuey, M. C. (2010). Synchronous-voice computer-mediated communication: Effects on pronunciation. *CALICO Journal*, *28*, 1–20.

Bueno-Alastuey, M. C. (2011). Perceived benefits and drawbacks of synchronous voice-based computer-mediated communication in the foreign language classroom. *Computer Assisted Language Learning*, *24*, 419–432.

Bueno-Alastuey, M. C. (2013). Interactional feedback in synchronous voice-based computer mediated communication: Effect of dyad. *System*, *41*, 543–559.

Chamberlin Quinlisk, C. (2008). Nonverbal communication, gesture, and second language classrooms: A review. In S. G. McCafferty & G. Stam (Eds.), *Gesture: Second language acquisition and classroom research* (pp. 25–44). New York, NY: Routledge.

Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2015). Second language comprehensibility revisited: Investigating the effects of learner background. *TESOL Quarterly*, *49*, 814–837.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Derwing, T. M., Diepenbrooke, L. G., & Foote, J. A. (2012). How well do general-skills ESL textbooks address pronunciation? *TESL Canada Journal*, *30*, 23–44.
Derwing, T. M., & Munro, M. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *30*, 379–397.
Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Philadelphia, PA: Benjamins.
Derwing, T. M., Munro, M. J., & Weibe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, *48*, 393–410.
Eckman, F. R. (2008). Typological markedness and second language phonology. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 95–115). Philadelphia, PA: Benjamins.
Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University Press.
Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, *51*, 281–318.
Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, *39*, 399–423.
Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-linguistic research* (pp. 233–277). Timonium, MD: York Press.
Foote, J., Holtby, A., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal*, *29*, 1–22.
Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of non-native speech. *Language Learning*, *34*, 65–89.
Gass, S., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory setting. *Language Learning*, *55*, 575–611.
Gass, S., & Mackey, A. (2015). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 180–206). New York: Routledge.
Gurzynski-Weiss, L., & Baralt, M. (2014). Exploring learner perception and use of task-based interactional feedback in FTF and CMC modes. *Studies in Second Language Acquisition*, *36*, 1–37.
Hardison, D. M. (1999). Bimodal speech perception by native and non-native speakers of English: Factors influencing the McGurk effect. *Language Learning*, *49* (Suppl. 1), 213–283.
Horgues, C., & Scheuer, S. (2014). "I understood you, but there was this pronunciation thing…": L2 pronunciation feedback in English/French tandem interactions. *Research in Language*, *12*, 145–161.
Isaacs, T., & Thomson, R. (2012). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly*, *10*, 135–159. doi: 10.1080/15434303.2013.769545
Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, *23*, 83–103.
Jepson, K. (2005). Conversations—and negotiated interaction—in text and voice chat rooms. *Language Learning & Technology*, *9*, 79–98.
Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *Modern Language Journal*, *94*, 554–566.
Keck, C., Iberri-Shea, N., Tracy-Venture, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Philadelphia, PA: John Benjamins.
Kennedy, S., Guénette, D., Murphy, J., & Allard, S. (2015). Le rôle de la prononciation dans l'intercompréhension entre locuteurs de français lingua franca [The role of pronunciation on comprehension between speakers of French as a lingua franca]. *Canadian Modern Language Review*, *71*, 1–25.
Lee, J., Jang, J., & Plonsky, L. (2014). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, *36*, 1–23.
Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*, 369–377.
Levis, J. (2016). Research into practice: How research appears in pronunciation teaching materials. *Language Teaching*, *49*, 423–437.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Lin, H. (2015). Computer-mediated communication (CMC) in L2 oral proficiency development: A meta-analysis. *ReCALL*, *27*, 261–287. doi: 10.1017/S095834401400041X
Loewen, S., & Philp, J. (2012). Instructed second language acquisition. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 53–73). Malden, MA: Blackwell Publishing.
Loewen, S., & Wolff, D. (2016). Peer interaction in F2F and CMC contexts. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 162–184). Amsterdam, The Netherlands: John Benjamins.
Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, *4*, 126–141.
Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, *21*, 557–587.
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. 407–449). Oxford, UK: Oxford University Press.
Matsumoto, Y. (2011). Successful ELF communications and implications for ELT: Sequential analysis of ELF pronunciation negotiation strategies. *Modern Language Journal*, *95*, 97–114.
Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*, 43–52.
Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*, 73–97.
Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*, 520–531.
Munro, M. J., & Derwing, T. M. (2015). A prospectus for pronunciation research in the 21st century: A point of view. *Journal of Second Language Pronunciation*, *1*, 11–42.
Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, *28*, 111–131.
Nakahama, Y., Tyler, A., & van Lier, L. (2001). Negotiation of meaning in conversational and information gap activities: A comparative discourse analysis. *TESOL Quarterly*, *35*, 377–405.
Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. New York: Routledge.
Norris, J. M. (2015). Statistical significance testing in second language research: Basic problems and suggestions for reform. *Language Learning*, *65* (Suppl. 1), 97–126.
Ortega-Llebaria, M., & Colantoni, L. (2014). L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in Second Language Acquisition*, *36*, 331–353.
Philp, J., Adams, R., & Iwashita, N. (2014). *Peer interaction and second language learning*. New York, NY: Routledge.
Pica, T. (1987). Second language acquisition, social interaction, and the classroom. *Applied Linguistics*, *8*, 3–21.
Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction and research. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9–34). Clevedon, UK: Multilingual Matters.
Pickering, L. (2009). Intonation as a pragmatic resource in ELF interaction. *Intercultural Pragmatics*, *6*, 235–255.
Plonsky, L. (2015). Statistical power, p values, descriptive statistics, and effect sizes: A "back-to-basics" approach to advancing quantitative methods in L2 research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 23–45). New York, NY: Routledge.
Saito, K. (2015). Communicative focus on second language phonetic form: Teaching Japanese learners to perceive and produce English /ɹ/ without explicit instruction. *Applied Psycholinguistics*, *36*, 377–409. doi: 10.1017/S0142716413000271

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning*, *62*, 595–633.
Saito, K., & Shintani, N. (2016). Do native speakers of North American and Singapore English differentially perceive comprehensibility in second language speech? *TESOL Quarterly*, *50*, 421–446. doi: 10.1002/tesq.234
Saito, K., Trofimovich, P., & Isaacs, T. (2015). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*. Online Advance Publication. doi: 10.1093/applin/amv047
Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, *37*, 217–240. doi: 10.1017/S0142716414000502
Sato, M., & Ballinger, S. (2016). Understanding peer interaction: Research synthesis and directions. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 1–30). Amsterdam, The Netherlands: Benjamins.
Sato, M., & Fujii, A. (2013). *International versus local computer-mediated peer interaction: An exploratory study*. Paper presented at the 5th Biennial International Conference on Task-Based Language Teaching. Edmonton, AB: University of Alberta.
Sauro, S. (2011). SCMC for SLA: A research synthesis. *CALICO Journal*, *28*, 369–391.
Sauro, S., & Smith, B. (2010). Investigating L2 performance in text chat. *Applied Linguistics*, *31*, 554–577.
Sicola, L. (2008). *"No, they won't 'just sound like each other'": NNS-NNS negotiated interaction and attention to phonological form on targeted L2 pronunciation tasks* (Unpublished doctoral thesis). Philadelphia, PA: University of Pennsylvania.
Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition*, *26*, 365–398.
Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*, 661–699.
Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, *82*, 320–337.
Thomas, M. (1994). Assessment of L2 proficiency in second language acquisition research. *Language Learning*, *44*, 307–336.
Williams, J. (1999). Learner-generated attention to form. *Language Learning*, *51*, 303–346.
Winke, P., Gass, S., & Myford, C. (2013). Raters' L2 background as a potential source of bias in rating oral performance. *Language Testing*, *30*, 231–252.
Yanguas, I. (2010). Oral computer-mediated interaction between L2 learners: It's about time! *Language Learning & Technology*, *14*, 72–93.
Zhao, S. Y., & Bitchener, J. (2007). Incidental focus on form in teacher-learner and learner-learner interactions. *System*, *35*, 431–447.
Ziegler, N. (2016a). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition*, *38*, 553–586. doi: 10.1017/S027226311500025X
Ziegler, N. (2016b). Taking technology to task: Technology-mediated TBLT, performance, and production. *Annual Review of Applied Linguistics*, *36*, 136–163.
Zielinski, B. (2015). The segmental/suprasegmental debate. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 397–410). Hoboken, NJ: John Wiley and Sons.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

APPENDIX A

SCHOLARSHIP TASK

Directions

In this task, you will discuss three different candidates for a scholarship. First read the background information about the scholarship, and then read the information about each of the candidates. You and your partner both have the same three candidates, but you each have different information about them. Tell your partner about the information that you have for each candidate. (Do not show them the information sheet.) Decide which candidate should receive the scholarship.

Scholarship Requirements

Your government places great values on education and has decided to send some young people overseas to America to be educated in a leading university. These young people will get a scholarship to pay for all their university fees and living expenses for four years.

The candidates do not need to speak fluent English. But, if a candidate doesn't speak fluent English, he/she must take an English course in America before starting their university study. This costs the government more money.

The purpose of the scholarship is to help young people in your country to understand the world better. The government wants young people to have experiences overseas, then come back home to help build a better society.

Please write down the name of the candidate that you think should get the scholarship.

Scholarship Winner: _____

Info Sheet (1) for Scholarship Candidates

Candidate 1: Albert

- Male.
- IQ of 110.
- Hard-working (studies eight hours every day).
- Hobbies: soccer, captain of the national high school soccer team.
- Visited America at the age of 15 as a soccer player.

Candidate 2: Gina

- Age: 20.
- Sometimes lazy at school.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

- Difficult childhood:
  - Parents killed in a car crash.
- Good at music.
  - Possibility of having lessons with a famous teacher in America.
- Uncle and aunt live in America.
- Current level of spoken English: upper-intermediate.

Candidate 3: Julia
- Female.
- IQ of 85.
- C grades for school subjects.
- Has a boyfriend in England.
- Her parents want her to study in America.
- Current level of spoken English: fluent.

Info Sheet (2) for Scholarship Candidates

Candidate 1: Albert
- Age: 18.
- A grades in every subject.
- Loves everything about America.
- Wants to be a professional soccer player, but plans to study business at the university.
- Current level of spoken English: pre-intermediate.

Candidate 2: Gina
- Female.
- IQ of 140.
- Usually A grades in school, but failed a class because she was partying before the final exam.
- Difficult childhood:
  - Stole a cell phone and was arrested at age 14. Was sent to a special school for troubled youth.
- Has never been to America.
- Wants to study painting and art history.

Candidate 3: Julia
- Age: 19
- Father is a diplomat.
  - Has traveled a great deal due to her father's work.
- Loves to help poor people:
  - Has won an award for her work with war refugees
- Wants to study International Relations.

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

APPENDIX B

My Favorite Things

In this activity, you will get a chance to get to know your partner better as the two of you ask questions about some of each other's favorite things. Take turns asking each other questions to find out the answers to the topics provided on the chart. Don't forget to ask your partner about the reasons behind their answers. Fill in the chart and write down the things that you have in common.

How many similarities are there?

Similarities
1.
2.
3.
4.
5.

My Favorite Things:

Favorite?	My Partner	Me
Book		
Movie		
Childhood memory		
Food		
Type of music/singer		
Pet		
Restaurant		
Actor		
Vacation location		
School subject		

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press

APPENDIX C

Transcription Key

Symbol	Feature
CAPITALIZATION	Emphasis
( )	Nonlinguistic information
(.)	Micropause
° °	Soft speech
?	Rising intonation
[ ]	Phonemic transcription
=	Linked speech
[	Overlapping speech
[	

https://doi.org/10.1017/S0272263116000449 Published online by Cambridge University Press