January 6, 2017

Task complexity, language-related episodes, and production of L2 Spanish vowels

Megan Solon, *University at Albany, SUNY*
Avizia Y. Long, *University of Guam*
Laura Gurzynski-Weiss, *Indiana University*

Studies in Second Language Acquisition

CAMBRIDGE UNIVERSITY PRESS

Task complexity, language-related episodes, and production of L2 Spanish vowels

Journal:	*Studies in Second Language Acquisition*
Manuscript ID	SLA-SI-2016-0040.R2
Manuscript Type:	Special Issue Article
Keywords:	task-based langauge teaching and learning, L2 pronunciation, task complexity, Cognition Hypothesis, L2 phonetics and phonology

SCHOLARONE™ Manuscripts

Running head: TASK COMPLEXITY, LREs, AND VOWEL PRODUCTION

Task complexity, language-related episodes, and production of L2 Spanish vowels

Megan Solon

University at Albany, SUNY

Avizia Y. Long

University of Guam

Laura Gurzynski-Weiss

Indiana University

Abstract

This study tests the theoretical predictions regarding effects of increasing task complexity (Robinson, 2001a, 2001b, 2007, 2010; Robinson & Gilabert, 2007) for second language (L2) pronunciation. Specifically, we examine whether more complex tasks (a) lead to greater incidence of pronunciation-focused language-related episodes (LREs) and (b) positively impact accuracy of phonetic form during task completion. Seventeen dyads of intermediate L2 Spanish learners completed simple (+few elements) and complex (-few elements) information-gap map tasks in which the pronunciation of Spanish vowels was made task essential through the inclusion of minimal pair street names (e.g., *Calle Copa* "Copa Street" and *Calle Capa* "Capa Street"). Results revealed no statistical difference in learner-produced pronunciation-related LREs in the simple and complex tasks. Vowel production, however, moved in a targetlike direction for one of five segments (/e/) during the complex task. Results therefore point to some benefits of task complexity manipulations for L2 pronunciation.

Task Complexity, Language-Related Episodes, and Production of L2 Spanish Vowels

A considerable amount of research on task-based language teaching and learning (TBLT) has examined the extent to which task characteristics and task sequencing, particularly manipulations of task complexity, can promote second language (L2) learning. The most utilized framework for this line of empirical study, the cognition hypothesis of L2 development (e.g., Robinson, 2001a, 2001b, 2007, 2010, 2011), and corresponding triadic componential framework (e.g., Robinson & Gilabert, 2007) and SSARC (simplify, stabilize-automatic-restructure, complexify) model (Robinson, 2010), operationalizes task complexity as the cognitive demands placed on learners (e.g., task complexity as cognitive complexity; for an alternative viewpoint on task complexity, see Skehan's limited capacity model; e.g., Skehan, 1996, 1998, 2009). Although the cognition hypothesis makes predictions regarding the effect of modifying cognitive complexity on learner production in terms of fluency, accuracy, and linguistic complexity presumably for all domains of language development, research to date has been limited to L2 grammar, lexis, and, most recently, pragmatics. Despite progress in understanding how manipulating task complexity may increase learning opportunities such as language-related episodes (LREs) and even L2 development for these aforementioned areas, we have yet to explore if and how the predictions of the cognition hypothesis apply for learners' attending to and development of L2 phonetics/phonology. Even though grammar-focused research has demonstrated that incidental focus-on-form episodes in task-based classrooms do include episodes focused on pronunciation and phonetic/phonological form (e.g., Ellis, Basturkmen, & Loewen, 2001; Loewen, 2005), it is still unknown whether manipulating task complexity can (a) positively impact accuracy of phonetic form (as it has been shown to do for grammatical form) during task completion and (b) lead to increased pronunciation-focused LREs, a specific learning opportunity considered to be an instantiation of learner reflection on language form that has been linked to linguistic development (Baralt, 2014).[1]

The present study extends research on the role of task complexity in TBLT and L2 development to the understudied realm of L2 pronunciation, focusing on the effects of manipulating task complexity in relation to the occurrence of pronunciation-focused LREs and the accuracy of English learners' L2 Spanish vowel production during dialogic face-to-face information gap tasks. This experiment provides initial evidence that specific task complexity manipulations (+/- few elements) may confer some benefits for improved accuracy of L2 pronunciation as they have often been shown to do for other aspects of L2 acquisition.[2]

Background

The Role of Attention in Task-Based Approaches to L2 Learning

Much like the larger field of L2 research (e.g., Schmidt, 1990, 2001), task-based researchers hold that learners must be encouraged to attend to form within meaning-based task interaction (Ellis, 2003; Long, 1985, 1998, 2015; Long & Norris, 2000; Robinson, 2001b, 2011; Skehan, 1998). The use of tasks is believed to mimic real-life language use and prepare learners for task completion outside of the L2 classroom. For this reason, within task-based methods, learners are directed to making form-meaning connections via theoretically based task design principles. Additionally, when learners notice the gap between their own interlanguage production and the target language during meaning-based interaction, they are encouraged to negotiate meaning, reformulate their hypotheses, and change their output production (Long, 1996; Swain, 1995, 2005).

One common manifestation of learner attention to form during meaning-based dyadic task-based interaction is the production of LREs, what Swain and Lapkin (2001) have operationalized as instances in which learners "talk about language they are producing, question their language use, or other- or self-correct their language production" (p. 104). As Baralt (2014) explains, LREs "are a conscious reflection on language that emerges from learners' collaborative task-based work…making them an effective way to operationalize learning opportunities" (p. 103). Importantly, and as pointed out by a reviewer, not all instances of reflection on form are captured in LREs. Thus, although they are a common measure appropriate for this dyadic task-based context, LREs, as with any methodological choice, are not without limitations. In this study, we examined the extent to which task complexity, as defined by Robinson (2001), influences the incidence of pronunciation-related LREs.

The Cognition Hypothesis and Task Complexity

The link between awareness of form within meaning-based interaction and L2 development is a central tenet of the cognition hypothesis (Robinson, 2001a, 2001b, 2007, 2010, 2011). The hypothesis and corresponding triadic componential framework (Robinson & Gilabert, 2007) and SSARC model (Robinson, 2010) offer a theoretical foundation for maximizing L2 learning through the manipulation of cognitive task demands and subsequent sequencing of tasks according to cognitive complexity. Theoretically, the hypothesis states that tasks should be designed and sequenced according to each task's cognitive complexity, or the "attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner" (Robinson, 2001b, p. 29). The basic theoretical premise of the cognition hypothesis is that learners should complete tasks that are increasingly more complex in order to foster L2 development and thereby prepare learners for real-world task performance. Increasing task complexity along certain task dimensions is hypothesized to bring about predictable changes both in learner production (moving from fluency to a focus on accuracy and complexity) and in interactional features (operationalized here as LREs), ultimately leading (at least in theory) to advances in L2 production (see review in Jackson & Suethanapornkul, 2013).

The triadic componential framework elucidates, among other things, specific task factors to be modified so that versions of tasks can be more or less cognitively demanding. For example, along resource-dispersing dimensions, tasks are designed to encourage learners to focus on fluency, or meaning; having planning time to complete a task (+planning time) is seen as less cognitively demanding (i.e., less complex) than not having planning time (-planning time). Along resource-directing dimensions, task modifications are designed to encourage a focus on accuracy and complexity, or form; learners who complete a task with few items (+few elements) have a less cognitively demanding task as compared to those who complete a task with many items (-few elements). Finally, the SSARC model specifies the sequence in which task versions (more or less complex) can be completed to maximize L2 learning opportunities (see also Baralt, Gilabert, & Robinson, 2014, for recent research on task sequencing). Together, the cognition hypothesis, triadic componential framework, and SSARC model specify how tasks may be designed and sequenced to encourage learners' awareness of certain linguistic targets within meaning-based task completion and, ultimately, to promote L2 development. Outcomes can be accounted for via general accuracy or fluency measures or, of particular relevance for the current study, accuracy measures of a specific L2 structure. Given the extensive range of studies investigating the predictions of the cognition hypothesis and the specific focus of the current study, we restrict our discussion to research that has investigated the effects of task complexity on incidence of LREs and/or accuracy development of specific L2 targets in the face-to-face mode.

Task Complexity and LREs

Studies testing the cognition hypothesis have often found relationships between task complexity and learner-produced LREs when task-based interaction has focused on specific grammar structures (Baralt, 2014; Kim, 2009, 2012; Révész, 2011) and pragmatic moves (Kim & Taguchi, 2015) in the face-to-face mode. For example, Baralt (2014) found that English-speaking learners of Spanish produced more LREs and demonstrated more development of the past subjunctive in face-to-face task sequences containing more complex tasks (+reasoning) than in those containing simpler tasks (-reasoning). Révész (2011) and Kim (2009, 2012) examined manipulations of +/- few elements as well as reasoning, finding higher LRE production in the more complex task conditions for English conjoined clauses (Révész) and question development (Kim).[3] Finally, Gilabert, Barón, and Llanes (2009) manipulated cognitive complexity (-here-and-now, +reasoning, and -few elements) in narratives, information-gap map, and decision-making tasks and found relationships between task complexity and learners' production of LREs focused on morphosyntax and lexis, among other interactional features, though results differed according to task type.

Most recently, Kim and Taguchi have examined the acquisition of pragmatics from the lens of the cognition hypothesis. In their 2015 investigation, they examined a potential link between task complexity (+reasoning demands), what they coined as pragmatic-related episodes ("PREs" ibid, p. 4), and development of request expressions for English L2 learners, finding that more complex tasks led to more PREs and sustained gains in knowledge of request-making on the delayed posttest.

Task Complexity and L2 Accuracy of Specific Targets

Studies have also examined potential links between task complexity and L2 accuracy of specific targets. For instance, Révész (2009) found a relationship between task complexity (-

here-and-now) and learners' development of English past progressive during interaction with a researcher in a laboratory setting. Baralt (2013) similarly found a link between task complexity (+reasoning) and learners' development of Spanish past subjunctive when learners interacted with a researcher in an experimental context in the face-to-face mode (this relationship was not found for computer-mediated interaction). Examining learner-learner task-based interaction, Kim and Tracy-Ventura (2011) and Kim (2012) found that more complex tasks (Kim & Tracy-Ventura: +reasoning; Kim: +reasoning and -few elements) led to greater development of the English past tense and English question formation, respectively. Baralt (2014) found learner-learner interaction to lead to greater LREs and greater accurate use of the Spanish past subjunctive when learners interacted in task sequences that were more complex (complex-complex-simple or complex-simple-complex) in a classroom setting. As mentioned earlier, Kim and Taguchi (2015) have extended this line of work to the development of L2 pragmatics, finding task complexity (+reasoning) to relate to L2 English learners' development of request expressions, at least on delayed posttests.

Thus, when reviewing the literature on task complexity and LREs and/or accuracy of specific L2 targets, there is considerable evidence that increasing task complexity can lead to learners' attending to (Baralt, 2014; Gilabert et al., 2009; Kim, 2009, 2012; Révész, 2011) and accurate use of grammatical structures (Baralt, 2013; Kim, 2012; Kim & Tracy-Ventura, 2011; Révész, 2009) and pragmatics (Kim & Taguchi, 2015). To date it is unexplored as to whether or not task complexity may offer similar benefits for L2 pronunciation.

Acquisition of L2 Pronunciation

Pronunciation—broadly conceived of as "the ways in which speakers use their articulatory apparatus to create speech" (Derwing & Munro, 2015, pp. 2)—is an area of L2 development in which learners' linguistic abilities are known to often not reach nativelike levels (Bongaerts, 1999; Moyer, 1999). The present study focuses specifically on one aspect of pronunciation—segmental production—and uses acoustic measures to account for patterns in and changes to the pronunciation of individual segments during task completion. Previous research has indicated that, despite well-established difficulties in ultimate attainment in phonology, acquisition of specific consonants and vowels appears possible even by adult learners (Flege, 1987; Flege, Takagi, & Mann, 1995), although results differ between specific segments and are influenced by factors such as learner individual differences, language pairings, and other linguistic factors such as lexical frequency (Munro & Derwing, 2008). Still, accuracy in segmental production has been shown to be important to learners' comprehensibility (Issacs & Trofimovich, 2012; Saito, Trofimovich, & Issacs, 2015) and accentedness (Saito et al., 2015).

Given previous support of the assertion of the cognition hypothesis that greater task complexity can encourage greater accuracy, in this study, we explore whether more complex tasks encourage greater segmental accuracy, operationalized here as the production of vowel segments with acoustic properties that are closer to the properties of native speakers' vowel productions. As previously reviewed, this hypothesis rests on the contention that greater task complexity encourages more attention to form. In pronunciation-related literature, attention to phonetic form—typically approached through the lens of explicit instruction, training, and/or feedback—has been shown to be an important factor in L2 pronunciation development. However, only sparingly and/or indirectly has task-based research specifically focused on learners' attention to, or awareness of, phonetic form.

The role of attention in L2 pronunciation development. Pronunciation-based intervention studies have shown that, in general, training, instruction, and/or corrective feedback positively facilitate learners' production of L2 sound segments. For instance, Bradlow, Pisoni, Akahane-Yamada, and Tohkura (1997) found that training Japanese listeners to identify particular English segments (i.e., /r/ and /l/) resulted in improvement in the subjects' production of these segments as rated by American English listeners. Similarly, Saito and Lyster (2012a, 2012b) showed that drawing learners' attention to pronunciation features through recasts facilitated pronunciation development with regard to the targetlike segmental production of English /ɹ/ and /æ/ by Japanese learners. More broadly, a recent meta-analysis by Lee, Jang, and Plonsky (2015) demonstrated that pronunciation instruction is effective and beneficial for learners across proficiency levels, and larger effect sizes were found for studies incorporating corrective feedback (given on the part of the treatment provider—a teacher, researcher, teacher-researcher, or computer) than for studies that did not include feedback. Taken together, these findings provide support for the notion that learners' attending to pronunciation encourages improvement in production and embolden TBLT researchers to investigate whether tasks could be used to encourage such attention to pronunciation within meaning-based interaction.

Within TBLT, however, relatively few studies have specifically examined attention to L2 pronunciation, although several studies have acknowledged and/or accounted for learners' attending to phonological form within the context of a larger study. For instance, 17% of episodes of negotiation between Bitchener's (2004) English as a second language (ESL) learner-learner dyads completing two different communicative tasks were focused around pronunciation, and learners modified nearly 70% of the pronunciation problems that were signaled. Similarly, García Mayo (2005) showed that learners both self- and other-correct pronunciation during task-based interaction with other learners. In her study, of 33 repair utterances, five focused on phonological form; three of these were self-corrected, and two were other-repaired. Finally, Bueno-Alastuey's (2013) investigation of synchronous voice-based computer-mediated communication between mixed proficiency ESL learner-learner and learner-native speaker dyads showed that LREs focused on phonetic form constituted an average of 40% of all LREs produced during the completion of a two-way information exchange task: 40% for same-first language (L1; Spanish) learner-learner dyads, 35% for different-L1 (Spanish-Turkish) learner-learner dyads, and 49% for learner-native speaker dyads.

To our knowledge, the only previous study to date to specifically focus on attention to phonological form during interactive tasks as defined, broadly speaking, within TBLT literature is Sicola (2008) who, in her investigation of intermediate-level ESL learners, used an interactive map task similar to the present study's to necessitate accurate production of the interdental fricative /θ/ (i.e., the "th" sound in *thought* or *threw*) for successful task completion. Sicola examined (a) whether learners drew each other's attention to phonological form via corrective feedback during interaction encouraged by this task, (b) whether learners modified their production of /θ/ as a result of corrective feedback, and (c) whether these modifications resulted in more targetlike /θ/ realizations. Sicola's results showed that the learners used various strategies, such as pausing or rising intonation, to make their production of /θ/ more salient to their partners, provided a variety of corrective feedback types to one another, and, overall, modified their production of /θ/ in a targetlike direction in a majority (i.e., 65%) of cases.

The present study continues this line of research by examining whether the manipulation of task factors—specifically task complexity as operationalized following the cognition hypothesis—impacts incidence of pronunciation-focused LREs and accuracy in phonetic production during task completion. We selected Spanish vowels (see following subsections) as the target form and seek to determine whether the manipulation of cognitive complexity in tasks designed to make segmental (vowel) accuracy essential for successful task completion encourages learner-generated pronunciation-related LREs and accuracy in the production of L2 Spanish vowels.

Vowels in English and Spanish. Spanish vowels were chosen as the target structure given that targetlike production of Spanish vowels is a notable pronunciation challenge facing English-speaking L2 learners of Spanish (Schwegler, Kempff, & Ameal-Guerra, 2010). The Spanish vowel system is comprised of five monophthong vowels: the high front vowel /i/, mid-front /e/, low central /a/, mid-back /o/, and high back /u/. Spanish vowels are generally considered to be quite stable in the sense that few dialectal differences in Spanish vocalic production have been noted (although see Willis, 2005, 2008) and little to no variation is expected between vowels in stressed and unstressed contexts (see however Cobb & Simonet, 2015; Menke & Face, 2010). General American English, in contrast, contains approximately 14-15 vowels (e.g., Giegerich, 1992; Ladefoged, 2006), and although Spanish and English share several vocalic phonemes (i.e., /i e o u/), no Spanish vowel is identical to any English vowel (Hualde, 2005). English /i/, for example, is produced in a lower position than Spanish /i/ and glides upward (Stockwell & Bowen, 1965), English /u/ is generally produced in a more fronted position than Spanish /u/ (Bradlow, 1995), and English /e/ and /o/ tend to diphthongize (e.g., *tame* [teɪm] or *tote* [toʊt]). Additionally, Spanish has one low central vowel (/a/), whereas English has two: /æ/ (e.g., *tat*) and /ɑ/ (e.g., *tot*). English vowels also tend to be longer in duration than Spanish vowels, and English is known for the systematic reduction and centralization of vowels in unstressed positions (i.e., to schwa, [ə]; Ladefoged, 2006). Finally, dialectal variation in English vowel production is widely documented. English-speaking learners of Spanish, then, not only need to acquire the reduced (as compared to English) vocalic inventory of Spanish but also the differing phonetic/acoustic properties of Spanish vowels that reflect differences in position (i.e., higher/lower, more fronted/backed), duration, and variation by linguistic context. In the present study, we focus on accuracy in terms of vowel quality (i.e., position) and focus our analysis on vowels produced in stressed contexts.

L2 Spanish vowels. Previous empirical work on L2 Spanish vowels has examined vowel quality, vowel quantity (i.e., duration), tendency to diphthongize, and differences in quality and quantity on the basis of stress context (e.g., Cobb & Simonet, 2015; Díaz & Simonet, 2015; Elliott, 1997; Menke, 2010; Menke & Face, 2010; Simoẽs, 1996). Given the focus of the present study, we restrict our review to those studies that explore development in L2 Spanish vowel quality and, more specifically, to those that employ acoustic measures to track production and development.

Using formant measurements (which index vowel height and frontness/backness) to quantify production, Menke and Face (2010) examined Spanish vowel production during reading by 60 English-speaking learners at various levels of study and found that lower level learners (i.e., learners in their fourth semester of study) demonstrated significant difficulty in producing targetlike Spanish vowels, whereas advanced learners (i.e., Spanish majors and PhD students) produced vowels with nativelike formant values. Specifically, for lower level learners, front vowels /i/ and /e/ were produced in a significantly lower and more backed position than those of the more advanced learner groups and the native speakers. Additionally, the back vowel /u/ was produced in a more fronted, Englishlike position by lower level learners and a more backed (Spanishlike) position by more advanced learners. Overall, formant values for the advanced learners were similar to values found for six native speakers, and the acoustic space for these learners' vowels overlapped the space found for the native speakers.

Menke's (2010) cross-sectional examination of the development of Spanish vowels by native English-speaking learners in one of two immersion settings, a one-way (foreign language) or two-way (bilingual) immersion program, revealed that differences between the vowel spaces of learners in the two-way program and the native speakers decreased over time, whereas differences between the vowel spaces of learners in the one-way program and the native speakers increased. Nevertheless, the vowel space of all learners in Menke's study was larger—especially on the high-low dimension—than the vowel space of the native speakers, again suggesting influence of the L1 (English) in the production of Spanish vowels.

Finally, Cobb and Simonet's (2015) analysis of vowel productions of 15 female Spanish speakers (five intermediate learners, five proficient learners, and five native speakers) revealed the greatest differences between learner levels in the production of /u/: The intermediate learners produced /u/ in a more fronted (i.e., more Englishlike) position, whereas advanced learners produced /u/ very far back in the vowel space.[4]

Taken together, studies of L2 Spanish vowel production suggest that English-speaking learners of Spanish tend to move toward more targetlike production of vowels as learner level or experience increases and that the back vowels, especially /u/, seem to present the most common and persistent problems for learners. The present study adopts similar acoustic measurement techniques to examine if and how manipulations of task complexity, previously shown to encourage learner attention to and accuracy in other linguistic domains, affect incidence of pronunciation-related LREs and accuracy of L2 Spanish vowel production during task completion.

Present Study

To investigate the impact of manipulating task complexity on L2 pronunciation, the present study was guided by the following research questions: In a dyadic information-gap map task designed to make accurate vowel pronunciation task essential, to what extent do cognitive task demands affect (1) occurrence of pronunciation-focused LREs and (2) accuracy in the pronunciation of vowels in L2 Spanish performance? Following predictions of the cognition hypothesis and considering results from previous research, our hypotheses were as follows: There will be (1) greater incidence of pronunciation-focused LREs in the more complex tasks and (2) greater accuracy in the pronunciation of Spanish vowels in the more complex task.

Method

Participants and research site. Participants were 34 students (23 females, 11 males; *M* age = 18.7 years, *SD* = 0.7 years) from four intact Spanish Grammar in Context courses (a third-year course that serves as a bridge between basic language courses and content courses [i.e., literature, culture, and linguistics courses taught in the target language]) at a large Midwestern university. Of these, six participants reported being native or heritage speakers of Spanish; their data were excluded from the present analysis as were the data from their interlocutors (i.e., four L2 learners). For four additional learners, only one task version (i.e., either simple or complex) was successfully recorded; their data were also excluded. The remaining 20 participants were all L1 English-speaking learners of Spanish. Responses to a background questionnaire revealed participants had been studying Spanish for an average of 6.5 years (Min = 3 years; Max = 15 years), and two had studied abroad (one student for 3 months, the other for 7 weeks).[5]

Procedure and materials. Data collection took place during regularly scheduled class time. Students chose a partner and were provided with the study materials (i.e., sets of maps) and instructions. During the first task, participants were told that they had been hired as tour guides in the Spanish city of Toledo and that their job started the next day. Unfortunately, the maps that the tour company had given them to prepare for the tour were incomplete: One partner in each pair received a map with the route of the tour indicated by arrows and with all tour stops numbered but with no names for any locations (Version A); the other partner of each pair received a map with names of all locations throughout the city but no tour route nor stop numbers (Version B; see Appendix A for examples of both versions). Participants had to "call" their partners to help each other complete the missing portions of their maps. After completing the first task, participants were told that the company had decided to give them an additional route but had made the same mistake again. Pairs had to follow the same procedure to complete their maps for the new route through a new part of the city; this time, the information given to each participant (i.e., Version A or Version B) was switched. The tasks were designed to make accuracy in the production of Spanish vowels more task essential through the use of minimal pair street names in which the only difference between nearby streets was the vowel in stressed position (e.g., *Calle C**o**pa* "Copa Street" and *Calle C**a**pa* "Capa Street").

One set of maps comprised the simple task (see Appendix A), whereas the other set of maps comprised the complex task (Appendix B). Task complexity was manipulated according to Robinson's (2001, 2007) resource-directing variable +/-few elements: The simple map had +few elements in that the route had fewer stops (i.e., seven total) and the map contained fewer extraneous elements such as extra streets or landmarks not included on the tour; the complex map can be characterized as having -few elements in that the route had more stops (i.e., 11 total), and the map contained many more streets and additional landmarks not on the tour. Révész (2014) signaled the importance of providing independent evidence of the validity of the manipulations of constructs such as cognitive complexity when investigating cognitive models of

TBLT. We follow Baralt's (2013) employment of participants' retrospective estimation of time on task (as well as a posttask rating) as an independent indication of cognitive load: During the study, actual time on task was measured, and, in a posttask questionnaire, participants judged the difficulty of each task (on a scale from 1 "very easy" to 5 "very difficult") and estimated their time spent on each task, providing independent measures of task complexity.[6] Table 1 presents these results. It should be noted that participants were not made aware prior to task performance that they would be asked to assess task duration. As Block, Hancock, and Zakay's (2010) meta-analysis showed, in retrospective assessments such as those utilized in the present study, greater cognitive load is associated with an increase in the subjective-to-objective duration ratio (also included in Table 1). Paired samples *t* tests confirmed a statistical difference between the simple and complex tasks with respect to actual time on task, $t(19) = -6.07$, $p < .001$, $d = -2.09$; estimated time on task, $t(19) = -6.19$, $p < .001$, $d = 1.94$; and learner rating of task difficulty, $t(19) = -6.53$, $p < .001$, $d = -1.64$, with large effect sizes for all comparisons using Plonsky and Oswald's (2014) SLA-specific scale. However, a paired ratio *t* test revealed no significant difference and a small effect size in the subjective-to-objective duration ratio between simple and complex tasks, $t(19) = -.41$, $p = .683$, $d = 0.15$

INSERT TABLE 1 ABOUT HERE

To eliminate the potential confound of order of tasks and task complexity, half of the dyads performed the simple task first, whereas the other half performed the complex task first; task version (i.e., A or B described in the previous paragraph) was also counterbalanced across dyads. During task completion, participants were recorded individually directly onto desktop computers using Sanako SLH-07 head-mounted microphones and the software Sanako Study (Sanako Corporation, 2013); recordings were sampled at 32 kHz. Following the completion of the two map tasks, participants completed the previously mentioned exit and background questionnaires. A visual summary of the study procedure is provided in Figure 1.

INSERT FIGURE 1 ABOUT HERE

Coding.

LRE coding. Recordings of the learner-learner interactions were transcribed and coded for occurrence and type of LRE. Table 2 presents the types of LREs encountered in the data, their operationalization, and an example from the data.

INSERT TABLE 2 ABOUT HERE

Acoustic analysis. To measure pronunciation accuracy with regard to the production of vowels in L2 Spanish, the present analysis examined the formant structure of the Spanish vowels produced during interaction. This analysis focused on the production of vowels in stressed position and was limited to those vowels occurring between occlusive/approximant consonants (e.g., /p t k b d g/) or between one of these consonants and a pause. Each vowel token was isolated, and first and second formant (F1 and F2, respectively) measurements were taken from the midpoint of each production using a script in Praat (Boersma & Weenink, 2014). The settings included a window length of 0.025 s for all tokens and five formants estimated under 5500 Hz for females and under 5000 Hz for males. The F1 (see Figure 2) provides information regarding the vowel's production along the height dimension (i.e., the higher the F1 value, the lower the production of the vowel within the vocal tract), whereas the F2 provides information regarding the production along a front-back dimension (i.e., the higher the F2 value, the more fronted the production of the vowel). Each token was also coded for vowel type (i.e., /i e a o u/).

INSERT FIGURE 2 ABOUT HERE

To account and control for potential formant structure differences based on varying vocal tract sizes between speakers (i.e., especially of different sexes), all vowel productions were normalized using the speaker-intrinsic, vowel-extrinsic Lobanov method within the NORM suite (Thomas & Kendall, 2012), which is a web-based interface to the vowels package for the statistical software R. The Lobanov method in NORM uses each unique speaker's raw vowel measurements to calculate a normalized formant measurement for each vowel based on the mean and standard deviation for that formant throughout the speaker's vowel system.

To be able to comment on the targetlikeness (i.e., accuracy) of vowel productions, data were also elicited from two female English-Spanish bilingual native Spanish speakers (one from Southern Spain, one from Argentina) who were instructors of Spanish from the same institutional context as the L2 learners. These native Spanish speakers were audio-recorded completing the same tasks as the learners, and their vowel productions (n = 101) underwent the same acoustic analysis and normalization process. The native speaker data are not included in the statistical analyses but rather are used to provide a visual bilingual native speaker baseline to aid in interpreting learner productions (i.e., whether differences observed in learner productions by task constitute movement in a nativelike direction).

Statistical analysis. To answer the first research question (i.e., the extent to which cognitive complexity affects learner production of pronunciation-focused LREs), simple descriptive statistics were calculated for each LRE type documented. Then, to control for the differences in time on task between the simple and complex tasks, a rate of LRE production by dyad per minute was calculated, and the mean rate of pronunciation-related LRE production between the simple and complex tasks was compared using a paired samples *t* test. To answer the second research question (i.e., the extent to which cognitive complexity affects accuracy in the pronunciation of vowels in L2 Spanish performance), descriptive statistics were calculated for the F1 and F2 measurements taken for each vowel production. Because observations were not independent (i.e., each participant produced several tokens of each vowel) and because participants contributed an unequal number of each vowel, separate linear mixed models were generated for each vowel to examine the relationship between vowel quality (i.e., F1 and F2; the dependent variables) and task complexity (simple vs. complex), treating participant and token as random intercepts. These models included task order as an additional fixed effect to determine whether vowel quality was influenced by the order in which tasks were completed (i.e., simple-complex vs. complex-simple). We acknowledge that running this number of distinct models increases the likelihood of a Type I error. However, we decided that this was favorable to just two models (i.e., one for F1 and one for F2) because it facilitates the comparison of vowel quality for each individual vowel across categories of each fixed effect (e.g., /i/ in simple vs. complex task, /e/ in simple vs. complex task, etc.) as opposed to a comparison of vowel quality across vowels and categories of each fixed effect (e.g., /i/ in relation to /e/ in simple vs. complex task, /i/ in relation to /a/ in simple vs. complex task, etc.).

Results

LREs. Participants' LRE productions during the simple and complex tasks were categorized as one of the eight LRE types listed and detailed in Table 2. Table 3 presents the findings for the number of each LRE type produced overall by learners during the simple and complex tasks; LRE types in the shaded columns refer to those categories that are related to pronunciation.

INSERT TABLE 3 ABOUT HERE

Overall, there was greater incidence of LREs produced during completion of the complex task (n = 155) than during completion of the simple task (n = 102). LREs were distributed similarly across LRE types during the simple and complex tasks, lexis-focused and pronunciation-focused LREs being the two most common LRE types produced during both; however, pronunciation-focused LREs were highest during the simple task and lexis-focused LREs were highest during the complex task.

Focusing specifically on pronunciation-related LREs (the shaded columns in Table 3), it will be recalled that a rate of LRE production (i.e., LRE/min) was calculated; this choice was made to reduce the influence of a potential confound of time on task and task complexity. Table 4 presents the mean rate of pronunciation-related LRE production during the simple and complex tasks.

INSERT TABLE 4 ABOUT HERE

Our first research question inquired as to whether task complexity would impact learners' production of pronunciation-related LREs. As can be seen in Table 4, pronunciation-related LREs were produced at a higher rate during the simple task (i.e., 0.70 LRE/min) than during the complex task (i.e., 0.46 LRE/min). A paired samples *t* test revealed that these differences were not significant and the effect size was relatively small, $t(11) = 1.73$, $p = .112$, $d = 0.52$.

Pronunciation accuracy. To examine whether accuracy of phonetic form was impacted by task complexity manipulations, the pronunciation of L2 Spanish vowels was investigated. As a reminder, the target phonetic context for vowels examined in this study was in stressed position between stop phonemes or between a stop phoneme and a pause. Vowels in target contexts (N = 1,079) were isolated and analyzed acoustically along two dimensions: height (F1) and frontness-backness (F2). A total of 139 vowel tokens were excluded from the analysis due to production phenomena that interfered with accurate formant measurements (e.g., creaky voice, background noise); thus, the final vowel data set consisted of 940 tokens. The findings presented in this section are based on normalized formant measurements.

Table 5 presents the findings for production of each vowel by task complexity condition.

INSERT TABLE 5 ABOUT HERE

The mean formant values presented in Table 5 reflect tongue position in the oral cavity. Along the high-low dimension (i.e., F1), a lower value indicates a higher tongue position, whereas a higher value indicates a lower tongue position. Along the front-back dimension (i.e., F2), a lower value indicates tongue position that is in a more backed position, whereas a higher value indicates tongue position that is more fronted in the oral cavity.

As presented in Table 5, learners' production of /a/, on average, is strikingly similar during the simple and complex tasks. Their production of /i/ is similarly comparable regardless of the complexity of the task. For /e/, learners' productions are similar across tasks on the high-low dimension (i.e., F1) but are more fronted (i.e., have a higher F2) during the simple task than during the complex task. For /o/, learners' productions are comparable with regard to F1 but differ with regard to F2: Productions of /o/ are more backed during the simple task than during the complex task. Finally, learners' productions of /u/ are slightly higher and more backed during the complex task than the simple task.

To statistically examine the relationship between vowel quality (i.e., F1 and F2) and task complexity, we used multilevel modeling. Two linear mixed effects analyses were performed for each of the five Spanish vowels (/i e a o u/) using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2016) in R (R Core Team, 2016)—the first with normalized F1 values as the dependent variable and the second with normalized F2 values as the dependent variable. For each model, task order (simple-complex or complex-simple) and task complexity (simple or complex) were included as fixed effects, and participant and token were included as random intercepts. Visual inspections of residual plots were conducted to confirm normality of residuals for each model. The findings for the models for each vowel with normalized F1 values as the dependent variable are presented in Table 6, and the findings for the models with normalized F2 values as the dependent variable are presented in Table 7.

INSERT TABLE 6 ABOUT HERE

INSERT TABLE 7 ABOUT HERE

Table 6 shows that, along the F1 (vertical or height) dimension, there were no significant effects for task order or task complexity for any vowel. Table 7 shows that, along the F2 (horizontal or front-back) dimension, learners produced /e/ with significantly higher normalized F2 values during the simple task than during the complex task, indicating a more fronted /e/ production on the simple task and a more backed production during the complex task.

Although the findings for /e/ along the F2 dimension indicate an effect of task complexity on vowel production, they do not tell us whether participants' productions of this vowel are more accurate or nativelike during the complex task than the simple task. It will be recalled that two bilingual native speakers of Spanish completed the map task to provide a native Spanish vowel baseline for comparison. Figure 3 provides a vowel chart or plot of the acoustic space for learners for the simple and complex tasks separately, as well as for the native speakers for the simple and complex tasks combined in order to interpret the direction of effect (i.e., more or less accurate or nativelike) of the difference in /e/ production between the simple and complex tasks.

INSERT FIGURE 3 ABOUT HERE

In a vowel plot like that depicted in Figure 3, vowels are arranged both vertically and horizontally to provide a visual reference point of tongue position in the oral cavity. The vertical (i.e., F1) dimension corresponds to tongue height, with low vowels positioned toward the bottom of the figure and high vowels positioned toward the top. The horizontal (i.e., F2) dimension corresponds to tongue frontness or backness, with front vowels positioned toward the left of the figure and back vowels positioned toward the right. As can be seen in Figure 3, learners' productions of /e/ were more backed during the complex task than the simple task, which constitutes movement in a nativelike direction (i.e., is more similar to the native speakers' average /e/ productions). It is also interesting to note that, although not significant, many of the other descriptive differences observed in vowel formant averages between the simple and the complex tasks (i.e., Table 5) represent movement toward more nativelike vowels in the complex task (e.g., for /i/ and /u/ along the F2 dimension and for /o/ along F1 and F2). It should be recognized, however, that much individual variation existed in vowel production. As the error bars in Figure 3 indicate, speakers' productions of the vowels varied widely; we will return to this issue in the Discussion.

Discussion

The present study aimed to investigate whether the predictions of the cognition hypothesis (e.g., Robinson 2001, 2007), especially with regard to task complexity manipulations and their effect on learner production and attention to form, extend to L2 pronunciation outcomes. The cognition hypothesis predicts that greater task complexity will encourage greater incidence of interactional features and, in turn, greater accuracy. Support for these predictions has been found with regard to grammatical (e.g., Baralt, 2013; Kim, 2012) and pragmatic (Kim & Taguchi, 2015) targets. The present study investigated these predictions with regard to pronunciation and, specifically, segmental (vowel) production, which was made more task essential by employing a map task that involved navigating streets with names comprised of vocalic minimal pairs.

To examine incidence of interactional features, LREs were counted and labeled according to focus, and a rate of pronunciation-related LRE production (i.e., LRE/min) was calculated. It was shown that pronunciation-related LREs were produced at a higher rate during the simple task than the complex task, although the difference in mean rates was not statistically significant and the effect of task complexity condition was small. Thus, with respect to the first research question, the predictions of the cognition hypothesis were not borne out: There were no statistical differences in incidence of interactional features (operationalized as LREs) in simple versus complex tasks.

These results contrast with those studies reviewed previously that have explored the cognition hypothesis's predictions via grammar-focused LREs, all of which found evidence supporting greater reflection on form during complex tasks than during simple tasks. Such differences in results could point to a fundamental difference between grammatical and pronunciation targets with regard to attention to form and the role of task complexity. For example, whereas grammatical targets often have specific forms that can many times be described in metalinguistic "rules," a phonetic target necessarily refers to a variable and gradient range of production possibilities (e.g., a variety of /a/ productions fall within a native speaker range) and requires precise (but typically subconscious) physical modification of articulators to produce. Additionally, although it has been shown that explicit instruction and training are beneficial to pronunciation learning, more communicative approaches to language teaching often fail to incorporate explicit pronunciation training into the foreign language classroom (Pennington & Richards, 1986); in fact, pronunciation training was not a part of the basic language curriculum of the institutional context examined in this study. Thus, learners may be less familiar with verbally reflecting on phonetic form than they are for other targets. In the future, alternative conceptualizations of attention to form should be used to explore whether task complexity may, in fact, encourage greater attention to phonetic form but in a manner that is not accounted for by LREs.

Given that we adopted LREs as instantiations of reflection on form (and despite not finding the hypothesized results for task complexity effects on learner reflection on pronunciation), we wondered whether the present dataset could provide specific evidence that, in fact, reflection on phonetic form during the completion of the present tasks actually led to modifications in a targetlike direction (as Sicola, 2008, found for modifications made after corrective feedback). Previous research has suggested that learner-learner dyads do provide corrective feedback regarding pronunciation to each other and that signaled pronunciation problems are more often modified than unmodified in learner-learner interaction (e.g., Bitchener, 2004; Sicola, 2008). Nevertheless, several studies have pointed to an advantage for learner dyads with mixed L1s (as opposed to shared L1s) such that negotiation regarding pronunciation is more likely to move toward a targetlike form and away from learners' respective L1s (e.g., Jenkins, 2000; Long & Porter, 1985; Sicola, 2008). Our participants all shared English as their L1, and the pronunciation targets under investigation are segments that largely exist (but with phonetic differences) in English. Thus, it would be reasonable to suspect that attention to (vowel) phonetic form during interaction could, in fact, lead to movement toward the shared L1 (English) form and away from the Spanish vowel target. To preliminarily explore whether attention to form as accounted for by LREs did, in fact, encourage targetlike pronunciation modification, we compared vowels produced during pronunciation-related LREs to all other vowels produced during the task-based interaction (regardless of task complexity). Figure 4 presents a plot of the vowel space that compares learners' productions of vowels during pronunciation-related LREs to those produced at other times (i.e., not during LREs). The native speakers' vowel averages are included as well as baseline reference points.

INSERT FIGURE 4 ABOUT HERE

As can be observed in Figure 4, in general, all vowels produced during pronunciation-related LREs are more similar to the native speakers' vowel averages than vowels produced in non-LRE contexts. These preliminary findings suggest that, despite sharing a L1, the negotiation between learner-learner pairs in our study and the resulting reflection on form during interaction encouraged targetlike modifications in L2 vowel production.

With regard to segmental accuracy and whether the predictions of the cognition hypothesis hold for L2 pronunciation targets (Research Question 2; conceived of as more nativelike vowel production measured acoustically), learners' pronunciation was found to be significantly more targetlike in the complex task for the vowel /e/ along the F2 or front-back dimension. That is, during the complex task, learners' /e/ productions were realized in a more backed, Spanishlike position than in the simple task. The pronunciation of other Spanish vowels (i.e., /i/, /o/, /u/) also appeared to trend toward more nativelike realizations (Figure 3) during the complex tasks, but these differences were not statistical. These results suggest that some of the predictions of the cognition hypothesis (i.e., greater accuracy in more complex tasks), shown frequently for accuracy on specific grammatical structures, may also extend to pronunciation, although not equally across all segments studied.

Previous studies have shown that recasts in L2 pronunciation can be effective in encouraging more targetlike production (and perception) of L2 sounds (e.g., Saito, 2013; Saito & Lyster, 2012a, 2012b). Interpreted in light of the present results, we suggest that perhaps tasks, learner negotiation, and (pronunciation) form-focused interaction during tasks also provide opportunities for pronunciation learning, or at least restructuring. The present study, of course, focused only on one specific acoustic correlate of vowel production (i.e., formant structure as an index for vowel quality); future research would benefit from the exploration of other aspects of vowel production (e.g., duration, English-speaking learners' tendency toward diphtongization) as well as other (e.g., perceptual or ratings-based) measures of accuracy.

It is a limitation of this study that few individual difference factors—especially that of L2 proficiency—were accounted for during data collection, especially considering the individual variation observed in Figure 3. It is well known that individual characteristics of interactional partners can affect the interactive moves and language produced during interaction (e.g., Bueno-Alastuey, 2012, 2013; Kim & McDonough, 2008; Leeser, 2004; Mackey, Oliver, & Leeman, 2003; Pica, Lincoln-Porter, Paninos, & Linnell, 1996; Porter, 1983; Sato & Lyster, 2007; Varonis & Gass, 1985). Nevertheless, the issue of L2 proficiency in this context is complicated by the fact that many common methods of measuring L2 proficiency for use in empirical research (e.g., cloze tests; Tremblay, 2011) do not account for pronunciation abilities, and it is well attested anecdotally and empirically that grammatical knowledge or proficiency does not always correspond with pronunciation abilities (see Solon, 2013, for a discussion of proficiency measurements for L2 phonetic/phonological research). It is interesting to note that the participant who made some of the most dramatic modifications in segmental accuracy between the simple and complex tasks had been studying Spanish for 6 years and was interacting with a learner who had been studying Spanish much longer than the group's average (i.e., 14 years as compared to the group average of 6.5 years). This example, though singular, offers support for the idea that interaction between mixed proficiency learner dyads, and subsequent linguistic outcomes, may differ in key ways from that of matched proficiency dyads and that learner proficiency level is an important factor to consider. Future research that explores the relationship between task factors, interactional features, and pronunciation should carefully consider and account for learner proficiency levels.

Although the present study offers initial evidence of positive effects of task complexity manipulations for segmental accuracy, our knowledge of the effect of tasks on learner pronunciation and of the relevance of the predictions of the cognition hypothesis to issues of phonetic/phonological performance and learning would also benefit from future studies that investigate other types of tasks, such as monologic tasks or tasks that require longer or different types of production (e.g., narration, decision-making; see Gilabert et al., 2009, regarding differences in the impact of task complexity manipulations on interactional moves by task type), that include different manipulations of task complexity, and/or that measure pronunciation performance in other ways, including other phonetic/acoustic measures or, for example, via comprehensibility or intelligibility ratings. Additionally, when examining learner attention to form, measurements of attention and cognitive processing can and should be triangulated (e.g., via stimulated recall or think-aloud protocols). Future studies may also want to address the issue of the effect of the interlocutor, as the differences in vowel accuracy and/or the effect of LREs may differ when learners complete tasks with native or heritage speakers or more or less advanced learners than with learners from their same proficiency level (e.g., Bowles, Toth, & Adams, 2014; Bueno-Alastuey, 2013; Mackey et al., 2003; Varonis & Gass, 1985, among many others). Finally, investigations into the effect of mode are also warranted, as it is possible that pronunciation targets interact differently with mode (i.e., face-to-face versus computer-mediated) than do grammatical targets or more general performance measures. For example, Bueno-Alastuey's (2013) examination of synchronous computer-mediated communication found rates of pronunciation-related LREs (i.e., 35-49% depending on the characteristics of the interlocutors) similar to those observed for the face-to-face interaction in the present study even though pronunciation was not a specific focus of the task completed by her participants. The present study, thus, sets the stage for much future work examining the relationships between task complexity, learner attention to form, and pronunciation.

Conclusion

Corroborating the predictions of the cognition hypothesis, previous research has demonstrated that the manipulation of task complexity can lead to increased reflection on form and accuracy, often explored with regard to specific grammatical targets. The present study extended the scope of research on the cognition hypothesis and examined whether such predictions extend to specific pronunciation targets. Investigating reflection on L2 pronunciation as well as the production of L2 Spanish vowels by native English-speaking learners during task completion, the results of the present study indicate that greater task complexity did not lead to a higher rate of production of pronunciation-related LREs but did encourage greater accuracy for one segment (/e/). The results of the present study make an important first step in connecting previous research on the role of task factors to L2 learning at the phonetic and phonological level.

Tables

Table 1

Mean Time on Task (Actual and Participant-Estimated) and Perceived Difficulty of Simple and Complex Tasks with Standard Deviations (in Parentheses)

Task	*M* time on task (in min)	Estimated time on task (in min)	Subj-to-obj duration ratio	Perceived difficulty (1-5)
Simple	6.5 (1.9)	6.6 (2.0)	1.02	2.3 (1.1)
Complex	10.9 (2.3)	11.5 (3.0)	1.06	3.9 (0.9)

Note. Subj = subjective; obj = objective.

Table 2

LRE Type, Operationalization, and Examples

Type of LRE	Operationalization	Examples from dataset
Grammatical	Discussion about/correction of grammatical features of TL (including tense, word form, sentence structure, plural, article use, prepositions, question formation, pronoun, etc.; follows Kim, 2013)	P23: *¿Dónde vamos después de eso? ¿Adónde?* "Where are we going after this? To where?"
Lexical/meaning	Discussion of vocabulary choice, translation, or meaning (different from Kim, 2013; among others)	P30: *Uh ok…pasamos a la izquierda y un poco a el sur. A de Calle Dedo hasta Calle Dado y a el sur hasta el um…cuchara?* "Uh ok…we go to the left and a little south. To Dedo Street up to Dado Street and south until the um…spoon?" P29: *Calle…* "…Street" P30: *Es Calle Dedo* "It's Dedo Street" P29: *Uh-huh.* "Uh-huh"

		P30: *...hasta Calle Dado y sur en Calle Dado. Cuando y es cuando Calle Dado splits* "…up to Dado Street and south on Dado Street. When and it's when Dado Street splits"
Spelling	Episodes in which it was clear that the learner had the correct word but was focusing simply on the spelling.	P26: *Quique Estube. Q-u-i-q-u-e…Estube.*
Lexical + Spelling	Episodes focusing on vocabulary or word choice that also include spelling	P47: *Quique? Eh um ah…como se escribe quique?* "Quique? Eh um ah…how do you spell 'quique'"? P48: *No recuerdo "q"..ka? Ka..okay, u-i-k-u-e* "I don't remember "q"..ka? Ka..okay, u-i-k-u-e."
Pronunciation	Episodes focusing on the phonetic form or realization of words.	P23: *Ok. Eso es el Mosquito de San Patrichio..Pa- patricio* "Ok. This is the mosque of Saint Patrichio..Pa- Patricio"

Lexical/Pronunciation	Episodes in which learners discuss the pronunciation of the word, but whether the focus is about pronunciation or about the word is unclear	P80: *okay y entonces um en número dos es en la Calle Gubo* "Ok, and then um on number two is on Gubo Street." P79: *Dudo?* "Dudo?" P80: *Gubo. Gu:bo* "Gubo. Gu:bo"
Lexical/Pronunciation + Spelling	Same as above but also includes spelling	P25: *Se llama Universidad de Todedo* "It's called Todedo University" P26: *De tor..?* P25: *Tod…Todedo* P26: *Dordeo?* P25: *Todedo* P26: *Dodedo, ok.* P25: *'To' con 't' y dos 'd's'* "'To' with a 't' and two 'd's'" P26: *Todedo, ok bueno.* "Todedo, ok got it.
Pronunciation + Spelling	Episodes focusing on the phonetic form or realization of words that also include	P29: *Ok. Um.. Museo de catedrático*. "Um… Museum of the *Catedrático*"

spelling	P30: *Cate..?* P29: *Catedrático.* P30: *Drat?* P29: *El 'a' tiene un acento.* "The 'a' has an accent" P30: *Oh oh* P29: *Catedrahhh- [laughs]* P30: *[laughs] Ok* P29: *Sí.* "Got it"

Note. TL = target language

Table 3

Overall Number, Percentage, and Type of LREs Produced during Completion of Simple and Complex Tasks

Task complexity	Grammar	Lexis	Spell	Lexis + Spell	Pron	Lexis/Pron	Lexis/Pron + Spell	Pron + Spell	Total
Simple	8	25	12	6	28	21	0	2	102
	(8%)	(25%)	(12%)	(6%)	(28%)	(21%)	(0%)	(2%)	(100%)
Complex	19	54	12	9	34	24	3	0	155
	(12%)	(35%)	(8%)	(6%)	(22%)	(16%)	(2%)	(0%)	(100%)
Total	27	79	24	15	62	45	3	2	257
	(11%)	(31%)	(9%)	(6%)	(24%)	(18%)	(1%)	(1%)	(100%)

Note. Pron = Pronunciation; Spell = Spelling. The shaded columns refer to LRE types that are related to pronunciation.

Table 4

Mean Rate (LRE/min by Dyad) and Standard Deviation (in Parentheses) of Pronunciation-related LREs Produced during Completion of Simple and Complex Tasks

Task complexity	Pron	Lexis/Pron	Lexis/Pron + Spell	Pron + Spell	Overall rate of pronunciation-related LREs
Simple	0.39 (0.43)	0.28 (0.25)	0	0.03 (0.07)	0.70 (0.58)
Complex	0.26 (0.21)	0.19 (0.16)	0.02 (0.04)	0.02 (0.04)	0.46 (0.27)
Total	0.32 (0.33)	0.23 (0.22)	0.01 (0.03)	0.01 (0.05)	0.58 (0.46)

Note. Pron = Pronunciation; Spell = Spelling; LRE = language-related episode.

Table 5

Descriptive Statistics of Overall Normalized F1 and F2 Values for Vowels by Task Type (Simple and Complex)

Vowel	Simple task					Complex task				
	n	F1		F2		*n*	F1		F2	
		M	*SD*	*M*	*SD*		*M*	*SD*	*M*	*SD*
/i/	63	-0.82	0.71	1.38	0.51	101	-0.86	0.68	1.33	0.58
/e/	77	-0.03	0.46	0.86	0.33	112	-0.03	0.43	0.73	0.36
/a/	92	1.37	0.70	-0.50	0.23	105	1.37	0.62	-0.49	0.37
/o/	80	0.12	0.50	-1.11	0.41	115	0.15	0.47	-0.95	0.43
/u/	67	-0.69	0.89	-0.25	0.86	128	-0.83	0.44	-0.45	0.67

Table 6

Summary of Linear Mixed-Effects Statistics for Models Predicting Normalized F1 Values

Vowel	*n*	Fixed Effects	ß	*SE*	*t*	*p*	Random effects	Variance	*SD*
/i/	164	Intercept	-0.64	0.19	-3.37	.004	Participant	0.30	0.60
		Task complexity (simple)	-0.01	0.10	-0.13	.896	Token	0.00	0.00
		Task order (simple-complex)	-0.23	0.29	-0.81	.433			
/e/	189	Intercept	0.03	0.08	0.37	.713	Participant	0.03	0.17
		Task complexity (simple)	-0.03	0.09	-0.36	.724	Token	0.01	0.12
		Task order (simple-complex)	-0.15	0.10	-1.45	.169			
/a/	197	Intercept	1.08	0.22	4.86	< .001	Participant	0.33	0.58
		Task complexity (simple)	-.014	0.10	-1.45	.149	Token	0.32	0.56
		Task order (simple-complex)	0.30	0.27	1.11	.280			
/o/	195	Intercept	0.14	0.09	1.68	.109	Participant	0.05	0.23
		Task complexity (simple)	-0.03	0.06	-0.43	.665	Token	0.00	0.00
		Task order (simple-complex)	-0.08	0.13	-0.66	.519			
/u/	195	Intercept	-0.82	0.14	-6.07	< .001	Participant	0.15	0.39

Task complexity (simple)	0.16	0.85	1.88	.062	Token	0.00	0.00
Task order (simple-complex)	-0.04	0.20	-0.22	.830			

Note. Reference categories indicated in parentheses.

Table 7

Summary of Linear Mixed-Effects Statistics for Models Predicting Normalized F2 Values

Vowel	*n*	Fixed Effects	ß	*SE*	*t*	*p*	Random effects	Variance	*SD*
/i/	164	Intercept	1.33	0.08	17.21	< .001	Participant	0.02	0.14
		Task complexity (simple)	0.04	0.09	0.48	.632	Token	0.00	0.00
		Task order (simple-complex)	0.12	0.11	0.18	.857			
/e/	189	Intercept	0.62	0.07	8.82	< .001	Participant	0.02	0.15
		Task complexity (simple)	0.17	0.07	2.45	.021	Token	0.01	0.12
		Task order (simple-complex)	0.18	0.08	2.07	.055			
/a/	197	Intercept	-0.51	0.12	-4.12	< .001	Participant	0.03	0.16
		Task complexity (simple)	-0.01	0.07	-0.15	.882	Token	0.21	0.46
		Task order (simple-complex)	0.11	0.08	1.29	.212			
/o/	195	Intercept	-0.88	0.13	-6.76	< .001	Participant	0.03	0.17
		Task complexity (simple)	-0.00	0.15	-0.03	.980	Token	0.17	0.42
		Task order (simple-complex)	-0.18	0.10	-1.85	.086			
/u/	195	Intercept	-0.09	0.19	-0.46	.647	Participant	0.11	0.34

Task complexity (simple)	.021	0.21	-0.39	.701	Token	0.23	0.48
Task order (simple-complex)	-0.07	0.18	-0.38	.712			

Note. Reference categories indicated in parentheses.

References

Baralt, M. (2013). The impact of cognitive complexity on feedback efficacy during online versus face-to-face interactive tasks. *Studies in Second Language Acquisition, 35,* 689-725.

Baralt, M. (2014). Task complexity and task sequencing in traditional versus online language classes. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 59-122). London: Bloomsbury.

Baralt, M., Gilabert, R., & Robinson, P. (Eds.). (2014). *Task sequencing and instructed second language learning*. London: Bloomsbury.

Bitchener, J. (2004). The relationship between the negotiation of meaning and language learning: A longitudinal study. *Language Awareness, 13,* 81-95.

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica, 134,* 330-343.

Boersma, P., & Weenink, D. (2014). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.85, Retrieved from http://www.praat.org/

Bongaerts, T. (1999). Ultimate attainment in foreign language pronunciation: The case of very advanced late language learners. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 133-159). Mahwah, NJ: Erlbaum.

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.

Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America, 97,* 1916-1924.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohjura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101,* 2299-2310.

Bueno-Alastuey, M. C. (2012). Focus on form and negotiation of meaning in synchronous voice-based computer mediated communication: Effect of dyad. *Procedia – Social and Behavioral Sciences, 34,* 39-44.

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

Cobb, K., & Simonet, M. (2015). Adult second language learning of Spanish vowels. *Hispania, 98,* 47-60.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: Benjamins.

Díaz, M., & Simonet, M. (2015). Second language acquisition of Spanish /e/ and /ei/ by native English speakers. *Hispania, 98,* 750-761.

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.

Elliott, R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania, 80,* 95-108.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15,* 47-65.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /r and

/l/ accurately. *Language and Speech, 38,* 25-56.

García Mayo, M. D. P. (2005). Interactional strategies for interlanguage communication: Do they provide evidence for attention to form? In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 383-405). New York: Mouton de Gruyter.

Giegerich, H. J. (1992). *English phonology: An introduction*. New York: Cambridge University Press

Gilabert, R., & Barón, J. (2013). The impact of increasing task complexity on L2 pragmatic moves. In A. Mackey & K. McDonough (Eds.), *Second language interaction in diverse educational settings* (pp. 45-69). Amsterdam: Benjamins.

Gilabert, R., Barón, J., & Llanes, À. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *IRAL, 47,* 367-395.

Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge, UK: Cambridge University Press.

Jackson, D., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning, 63,* 330-367.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford, UK: Oxford University Press.

Kim, Y. (2009). The effects of task complexity on learner-learner interaction. *System, 37,* 254-268.

Kim, Y. (2012). Task complexity, learning opportunities, and Korean EFL learners' question development. *Studies in Second Language Acquisition, 34,* 627-658.

Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research, 12,* 211-234.

Kim, Y., & Taguchi, N. (2015). Promoting task-based pragmatics instruction in EFL classroom contexts: The role of task complexity. *Modern Language Journal, 99,* 656-677.

Kim, Y., & Tracy-Ventura, N. (2011). Task complexity, language anxiety and the development of past tense. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 287-306). Amsterdam: Benjamins.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in liner mixed effects models (Version 2.0-32). Available https://cran.r-project.org/web/packages/lmerTest/index.html

Ladefoged, P. (2006). *A course in phonetics* (5th ed.). Boston: Thomson.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36,* 345-366.

Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research, 8,* 55-81.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39,* 369-377.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition, 27,* 361-386.

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77-99). Clevedon, UK: Multilingual Matters.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bjatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.

Long, M. H. (1998). Focus on form in task-based language teaching. *University of Hawai'i Working Papers in ESL, 16*(2), 35-49.

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Malden, MA; Wiley.

Long, M. H., & Norris, J. M. (2000). Task-based teaching and assessment. In M. Byran (Ed.), *Encyclopedia of language teaching* (pp. 597-603). London: Routledge.

Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly, 19,* 207-228.

Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS–NNS and NNS–NNS adult and child dyads. *Language Learning, 53,* 35-66.

Menke, M. R. (2010). *The acquisition of Spanish vowels by native English-speaking students in Spanish immersion programs* (Unpublished doctoral dissertation). University of Minnesota, Minneapolis.

Menke, M. R., & Face, T., L. (2010). Second language Spanish vowel production: An acoustic analysis. *Studies in Hispanic and Lusophone Linguistics, 3,* 181-214.

Moyer, A. (1999) Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition, 21,* 81-108.

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning, 58,* 479-502.

Nuevo, A.-M. (2006). *Task complexity and interaction: L2 learning opportunities and interaction* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Nuevo, A.-M., Adams, R., & Ross-Feldman, L. (2011). Task complexity, modified output, and L2 development in learner-learner interaction. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 175-2020). Amsterdam: Benjamins.

Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly, 30,* 59-84.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64,* 878-912.

Porter, P. A. (1983). Variations in the conversations of adult learners of English as a function of the proficiency level of the participants (Unpublished doctoral dissertation). Stanford University, Stanford, CA.

R Core Team (2016). R: A language and environment for statistical computing (Version 3.2.4) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Available https://www.R-project.org/

Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition, 31,* 437-470.

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom□based study. *Modern Language Journal*, *95,* 162-181.

Révész, A. (2014). Towards a fuller assessment of cognitive models of task-based learning: Investigating task-generated cognitive demands and processes. *Applied Linguistics, 35,* 87-92.

Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for investigating task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). New York: Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22,* 27-57.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *IRAL, 45,* 193-213.

Robinson, P. (2010). Situating and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In M. Pütz & L. Sicola (Eds.), *Cognitive processing in second language acquisition* (pp. 243-268). Amsterdam: Benjamins.

Robinson, P. (2011). Task□based language learning: A review of issues. *Language Learning*, *61*(Suppl. 1), 1-36.

Robinson, P., & Gilbert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *IRAL, 45,* 161-176.

Saito, K. (2013). The acquisition value of recasts in instructed second language speech learning: Teaching the perception and production of English /ɹ/ to adult Japanese learners. *Language Learning, 63,* 499-529.

Saito, K., & Lyster, R. (2012a). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning, 62,* 595-633.

Saito, K., & Lyster, R. (2012b). Investigating the pedagogical potential of recasts for L2 vowel acquisition. *TESOL Quarterly, 46,* 387-398.

Sanako Corporation. (2013). Sanako Study Lab 1200 (Version 6.10) [Computer software]. Turku, Finland.

Sato, M., & Lyster, R. (2007). Modified output of Japanese EFL learners: Variable effects of interlocutor vs. feedback types. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 123-142). Oxford, UK: Oxford University Press.

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

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge, UK: Cambridge University Press.

Schwegler, A., Kempff, J., & Ameal-Guerra, A. (2010). *Fonética y fonología españolas* (4th ed.) Hoboken, NJ: Wiley.

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 dissertation). University of Pennsylvania, Philadelphia.

Simoẽs, A. R. M. (1996). Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish. *Hispania, 79,* 87-95.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17,* 38-62.

Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics, 18,* 268-286.

Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency and lexis. *Applied Linguistics, 28,* 510-532.

Solon, M. (2013, October). *Measuring proficiency in cross-sectional L2 phonetic* production *studies: An examination of three methods*. Paper presented at the Second Language Research Forum, Provo, UT.

Stockwell, R. P., & Bowen, D. (1965). *The sounds of English and Spanish*. Chicago: The University of Chicago 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, UK: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-483). New York: Routledge.

Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 99-118). New York: Routledge.

Thomas, E. R., & Kendall, T. (2012). NORM (Version 1.1) [Software]. Retrieved from http://ncslaap.lib.ncsu.edu/tools/norm/biblio1.php

Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Studies in Second Language Acquisition, 33,* 339-372.

Varonis, E. M., & Gass, S. (1985). Non-native/Non-native conversations: A model for negotiation of meaning. *Applied Linguistics, 6,* 71-90.

Willis, E. W. (2005). An initial examination of Southwest Spanish vowels. *Southwest Journal of Linguistics, 24,* 185-198.

Willis, E. W. (2008). No se comen, pero sí se mascan: Variación de las vocales plenas en la República Dominicana. *Actas del XV Congreso Internacional de la Asociación de Lingüística y Filología de América Latina (ALFAL)*, Montevideo, Uruguay. 6pp.

Notes

1. LREs have also been referred to as learners' attempts to "draw attention to L2 form-meaning connections in the context of meaning-based communication" (Révész, 2011, p. 165).

2. Our review of studies that examined learner LRE production in the face-to-face mode and/or development in the accuracy of specific linguistic targets revealed general support for the cognition hypothesis and its assertions about the effect of task complexity. Nevertheless, it should be noted that studies examining other aspects of face-to-face interaction such as recasts, confirmation checks, and clarification requests, among others, have found mixed results with respect to the relationship between task complexity and amount of interactional features (e.g., Gilabert, Barón, & Llanes, 2009; Nuevo, 2006; Nuevo, Adams, & Ross-Feldman, 2011; Robinson, 2001b) as well as L2 development (e.g., Nuevo, 2006). The reader is directed to the recent scoping review of Jackson and Suethanapornkul (2013).

3. Révész (2011) also examined metalinguistic talk.

4. These three studies (Menke, 2010; Menke & Face, 2010; and Cobb & Simonet, 2015) also examine the effects of stress, especially with regard to the potential (more Englishlike) centralization of unstressed vowels. Because the present study examines only stressed vowels, the results regarding centralization are not reviewed here.

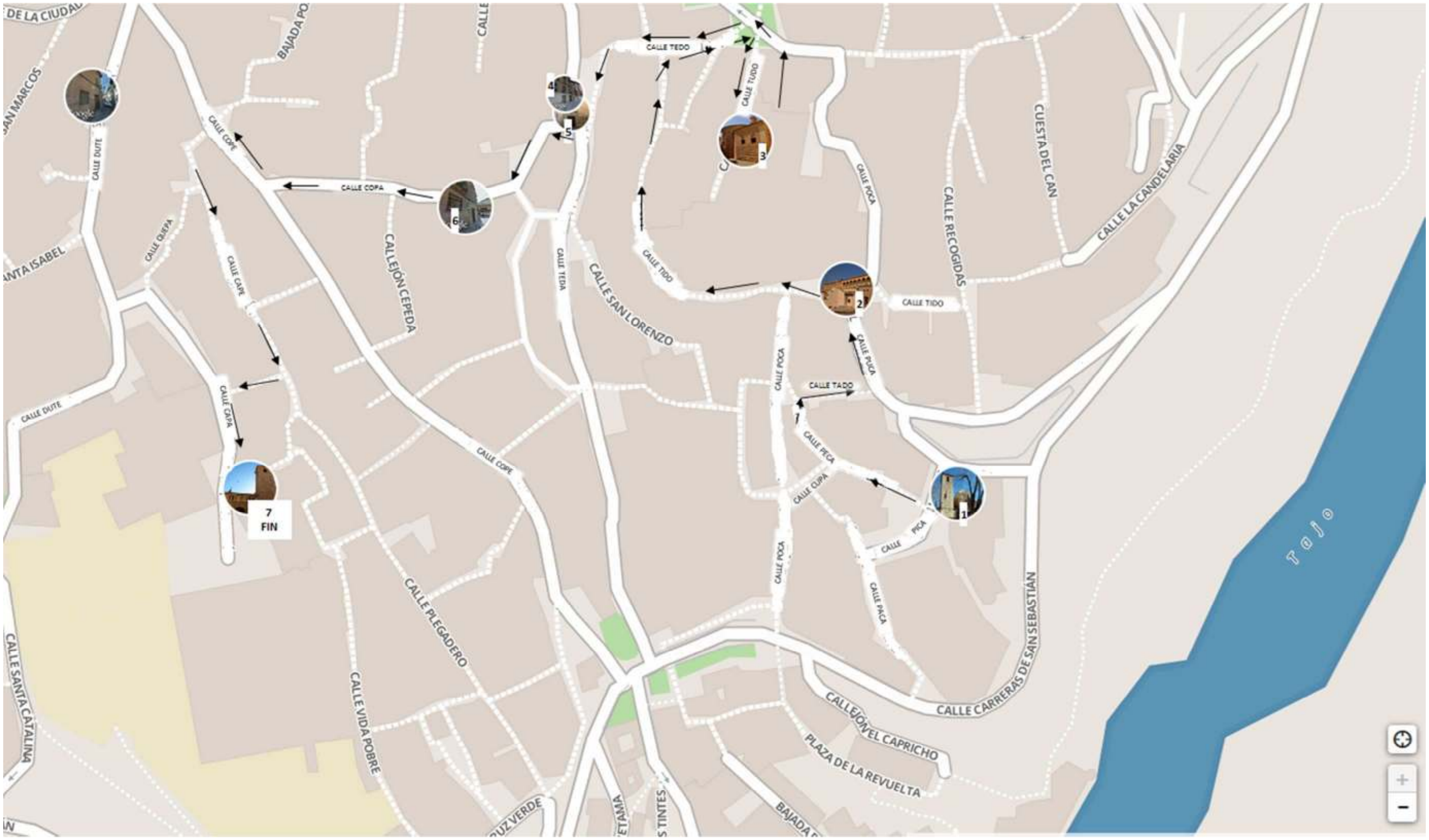
5. Although all participants came from the same university course level, it is a limitation of this study that no independent measure of proficiency was administered. The linguistic proficiency of interlocutors has been shown to be an influential factor in L2 interaction (e.g., Kim & McDonough, 2008; Leeser, 2004; Porter, 1983)—an issue that we return to in the Discussion section.

6. It is a limitation of this study that the time estimates and task difficulty ratings were performed after the completion of both tasks. These estimates and ratings would have been more valid and consistent with previous task complexity research had they been administered immediately after each task performance.

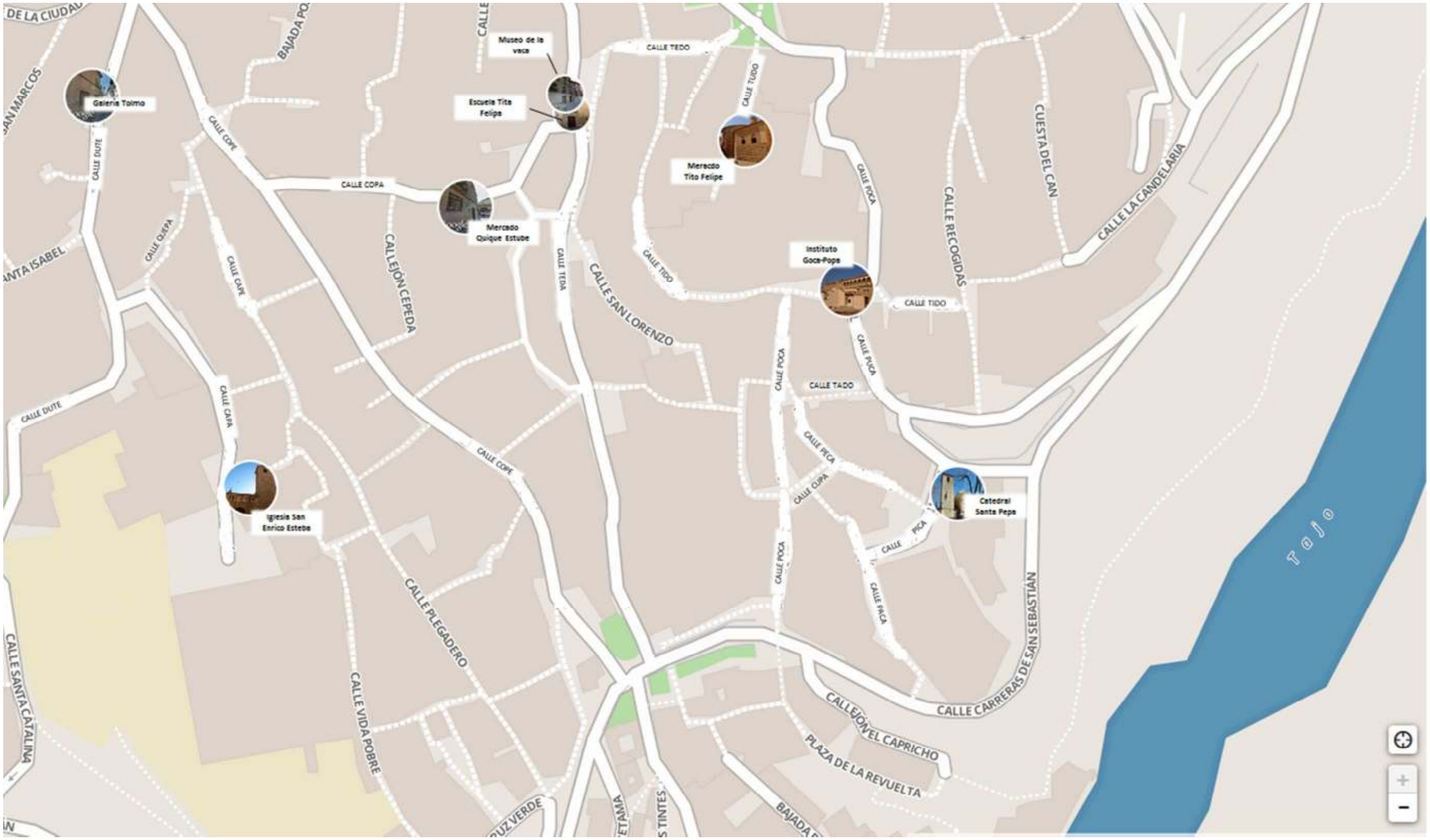
Appendix A

Sample maps (Simple version)

Version A



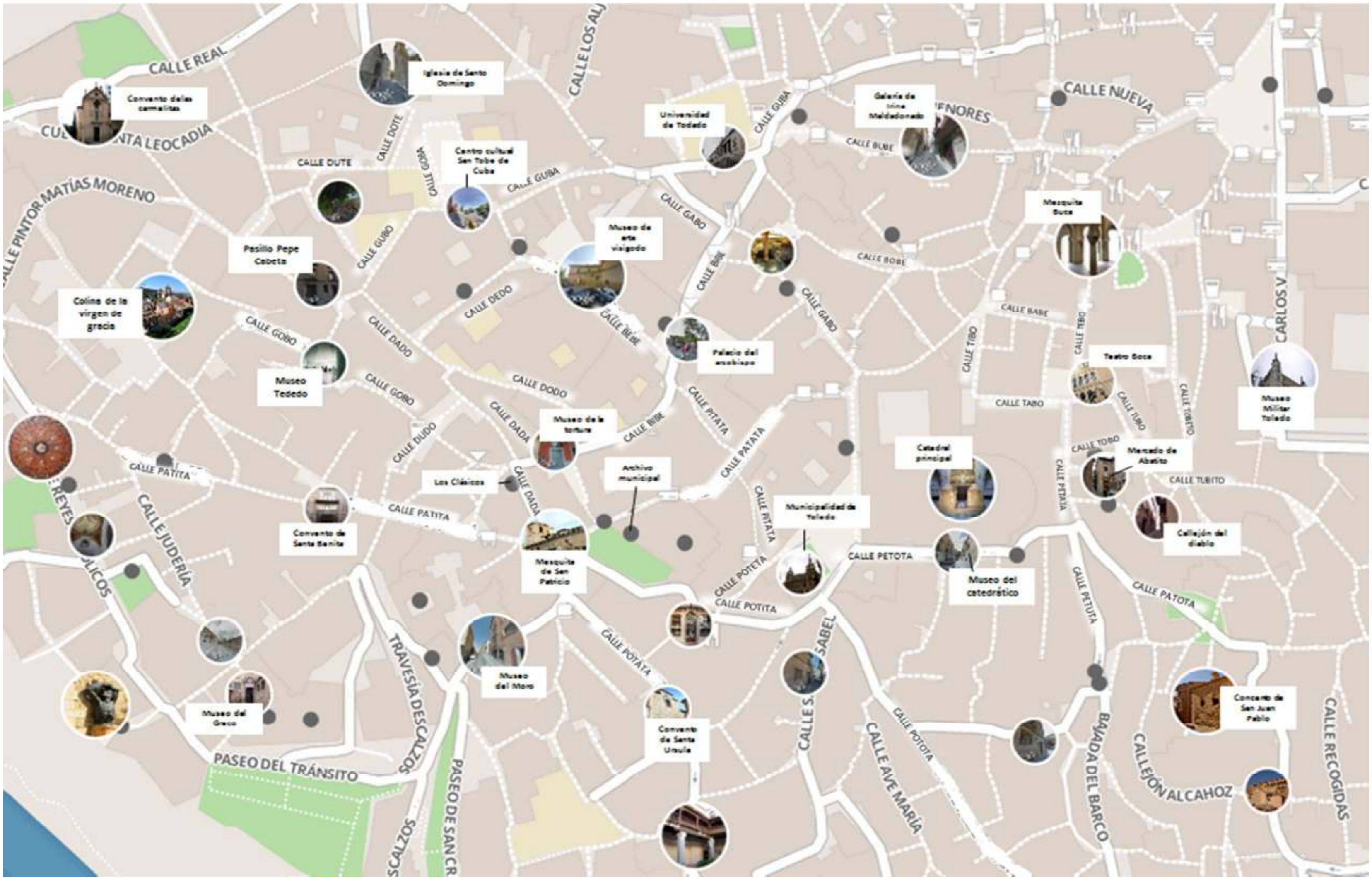
Version B


Galeria Tolmo
Museo de la vaca
Escuela Tito Felipe
Mercado Tito Felipe
Mercado Quique Estube
Instituto Goce-Pope
Iglesia San Enrico Esteba
Catedral Santa Pepa
CALLE TEDO
CALLE TUDO
CALLE COPA
CALLE COPE
CALLE DUTE
CALLE QUEPA
CALLE CAPE
CALLE CAPA
CALLEJÓN CEPEDA
CALLE TEDA
CALLE SAN LORENZO
CALLE TIDO
CALLE POCA
CALLE PUCA
CALLE TADO
CALLE PECA
CALLE CUPA
CALLE PICA
CALLE PACA
CALLE RECOGIDAS
CUESTA DEL CAN
CALLE LA CANDELARIA
CALLE CARRERAS DE SAN SEBASTIÁN
CALLEJÓN VEL CAPRICHO
PLAZA DE LA REVUELTA
CALLE PLEGADERO
CALLE VIDA POBRE
CALLE SANTA CATALINA
BAJADA PO
DE LA CIUDAD
SAN MARCOS
SANTA ISABEL
Tajo

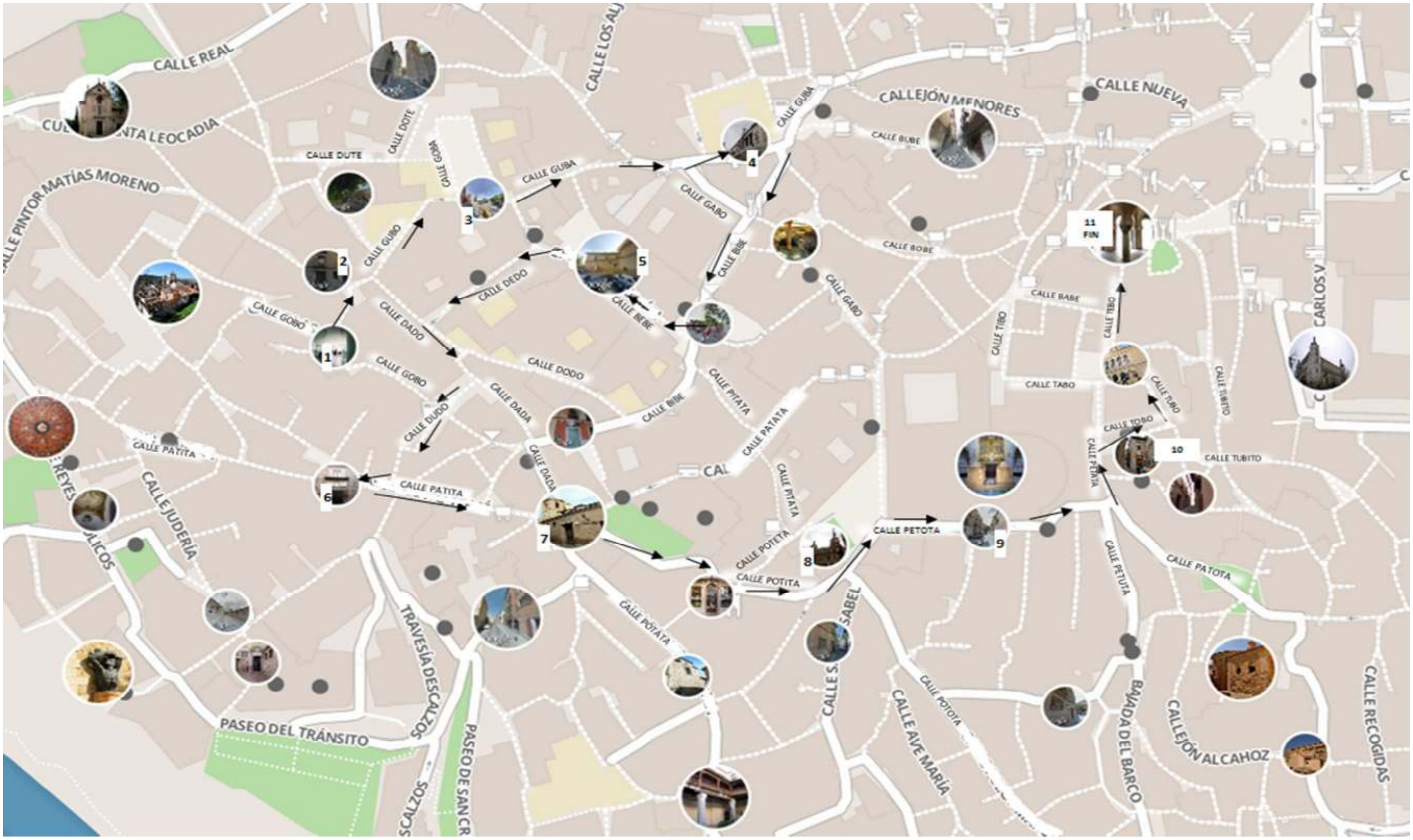
Appendix B

Sample maps (Complex version)

Version A



Version B


CALLE REAL
CALLE LOS ALJ
CALLEJÓN MENORES
CALLE NUEVA
CALLE DUTE
CALLE DOTE
CALLE GUBA
CALLE GABO
CALLE BUBE
CALLE BOBE
CALLE PINTOR MATÍAS MORENO
CALLE GUBO
CALLE GOBO
CALLE DADO
CALLE DEDO
CALLE BIBE
CALLE BEBE
CALLE BABE
CALLE TIBO
CALLE TABO
CALLE TUBO
CALLE TOBO
CALLE TUBITO
CALLE DODO
CALLE GOBO
CALLE DUDO
CALLE DADA
CALLE PITATA
CALLE PATATA
CALLE PATITA
CALLE JUDERÍA
CALLE DIDA
CALLE PITATA
CALLE POTETA
CALLE POTITA
CALLE PETOTA
CALLE PETATA
CALLE PETUTA
CALLE PATOTA
CALLE POTATA
CALLE POTOTA
CALLE AVE MARÍA
BAJADA DEL BARCO
CALLEJÓN ALCAHOZ
CALLE RECOGIDAS
CARLOS V
PASEO DEL TRÁNSITO
TRAVESÍA DESCALZOS
PASEO DE SAN CR
1
2
3
4
5
6
7
8
9
10
11 FIN

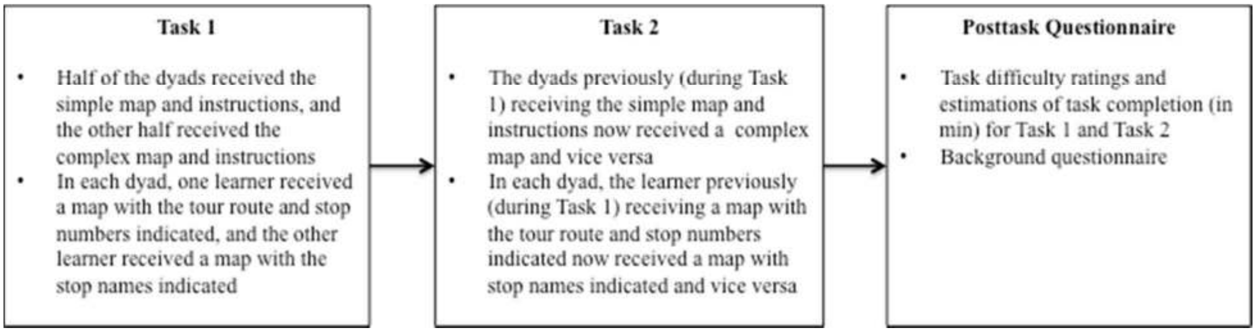


Figure 1. Visual schematic of study procedure.

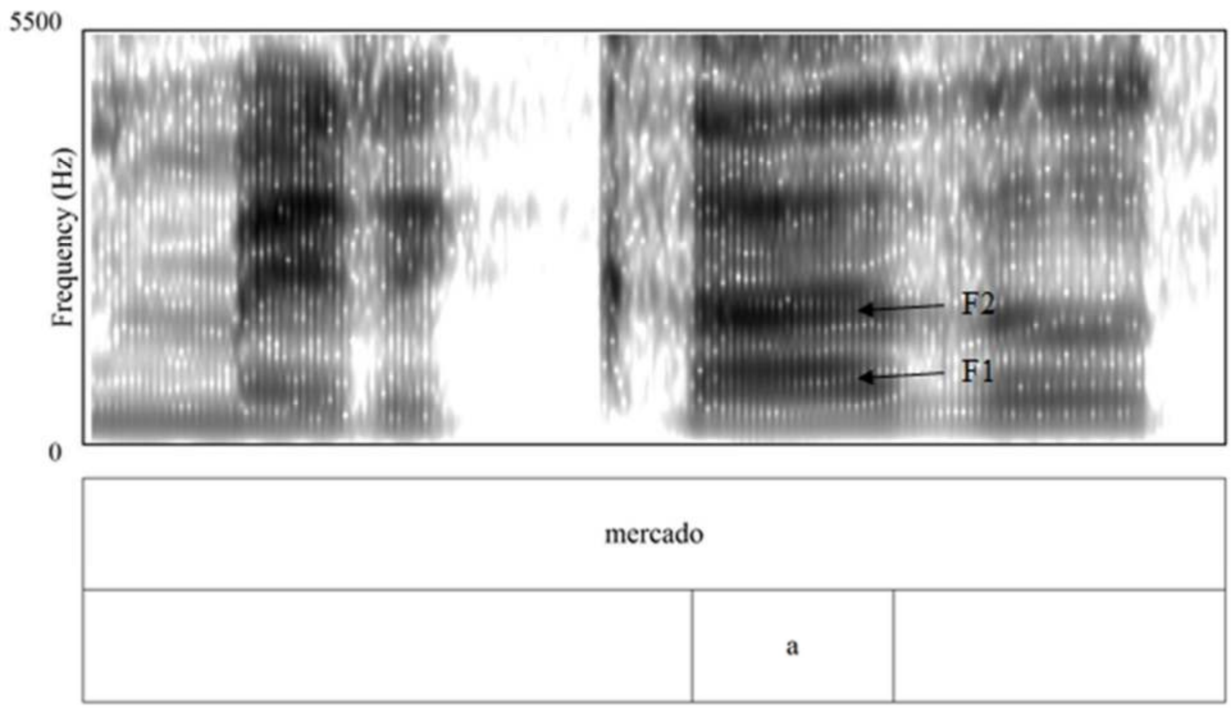


Figure 2. Production of Spanish word *mercado* “market” by one of the Spanish native speakers; figure shows the first and second formants (F1 and F2)—which index vowel height and frontness/backness, respectively—for the stressed vowel /a/.

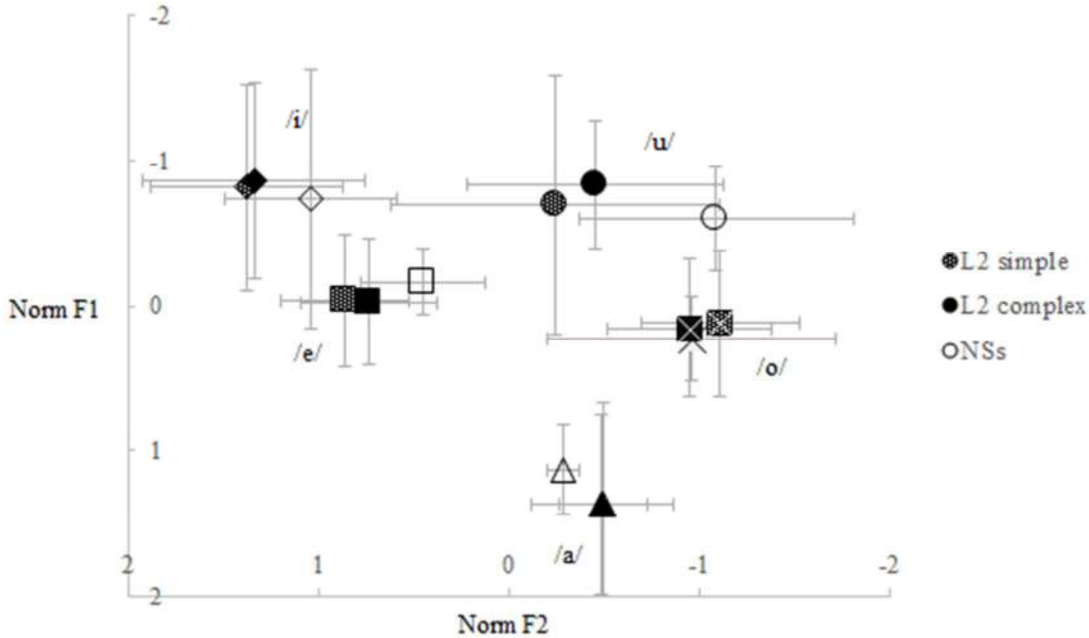


Figure 3. Plot of vowel space for L2 learners during simple task, L2 learners during complex task, and native speakers (NSs; simple and complex tasks combined).

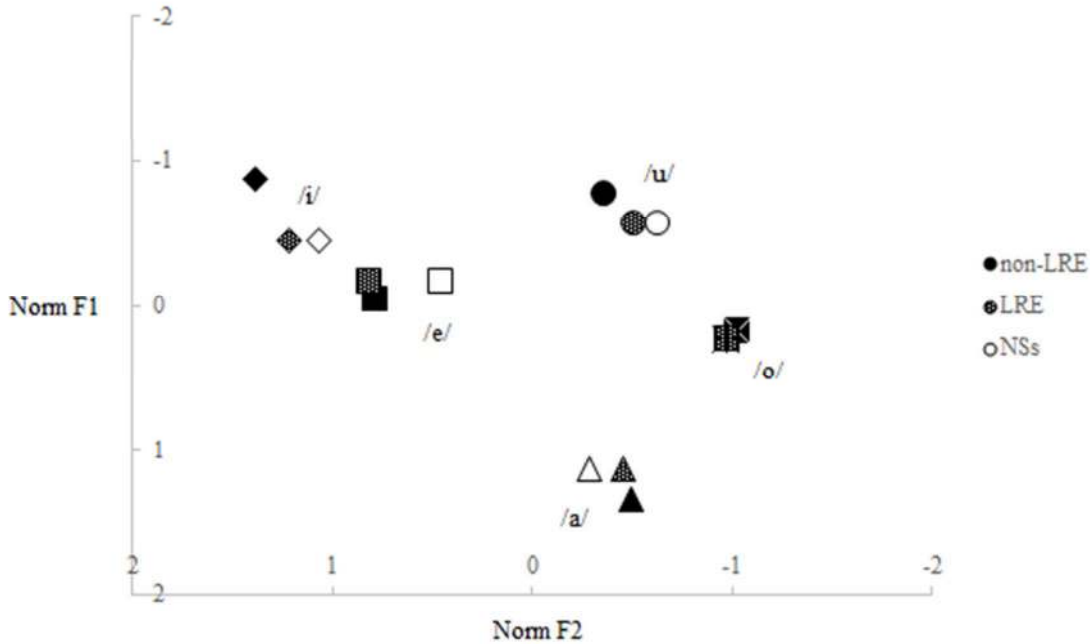


Figure 4. Vowel plot for L2 learners' productions in the context of a LRE and in a non-LRE context, as compared to native speaker (NS) vowel productions.