

A Reexamination of Ultimate Attainment in L2 Phonology: Length of Immersion, Motivation, and Phonological Short-Term Memory

Charles Nagle
Georgetown University

1. Literature review

1.1. *Factors implicated in second language pronunciation accuracy*

Research on the acquisition of second language (L2) phonology has identified a range of factors related to pronunciation accuracy: language immersion (Elliot 1995; Moyer, 1999; Shively, 2008), motivation (Purcell & Suter, 1980; Elliot, 1995a; 1995b; Moyer, 1999; Muñoz & Singleton, 2007; Shively, 2008), instruction (Elliot 1995b; Helmke & So, 1980; Moyer, 1999; Shively, 2008), phonological short-term memory (Aliaga-García, Mora, & Cerviño-Povedano, 2011; Cerviño-Povedano & Mora, 2010; MacKay, Meador, & Flege, 2001), and aptitude (Abrahamsson & Hyltenstam, 2008), just to name a few. Despite this growing body of research, results remain inconclusive due to methodological differences between studies. Researchers vary in their use of spontaneous or scripted tasks—tasks where participants speak extemporaneously or read a scripted paragraph or word set—and rating procedures. Participant samples are equally diverse both in terms of learners' proficiency levels and learning contexts.

Research on adults who have learned an L2 in a formal classroom setting has consistently shown that age, motivation, and instruction are important factors in L2 phonological development (Elliot, 1995a; 1995b; Moyer, 1999; Shively, 2008). In Elliot (1995a), intermediate-level adult learners of L2 Spanish completed four oral tasks: two oral mimicry tasks at the word and sentence level, a scripted word task, and a spontaneous speaking task. In addition to the dependent measures, participants took the Pronunciation Attitude Inventory (PAI), a Likert-type assessment where they rated statements as characteristically true of them or not. Judges rated participants samples on a scale from one to three (1 = *incorrect target sound*, 3 = *correct target sound*) on the oral mimicry and scripted word tasks and on a scale of one to five (1 = *very poor accent*, 5 = *native*) on the spontaneous speech task. A positive or negative attitude towards L2 pronunciation as assessed on the PAI was the most significant predictor of pronunciation accuracy for the individual tasks except word mimicry and globally across the four tasks. Elliot also demonstrated that attitude was significantly related to accuracy in a subsequent study on the role of formal instruction in L2 pronunciation development (1995b). Both Moyer (1999) and Muñoz and Singleton (2007) corroborate Elliot's findings with highly advanced learners who have near-native proficiency in the target language. Participants in Moyer (1999) were graduate teaching assistants and learners of L2 German; as in Elliot (1995a), they completed a series of speaking tasks at the word and paragraph level. Native judges later rated the samples as either native or nonnative and gave a confidence rating. Professional motivation accounted for 41% of the variance in the ratings and was the best predictor of pronunciation accuracy in a regression analysis (Moyer, 1999). Muñoz and Singleton (2007) obtained similar findings in their analysis of adult learners of L2 English; two of the

* I would like to thank Alfonso Morales-Front, Cristina Sanz, the audience of SLRF 2011, and the reviewers for their insightful suggestions on previous versions of this paper.

twelve were rated as native-like on a short speaking task (mean syllable length per sample = 9.9) and the researchers determined in interviews with the learners that each possessed high integrative or intrinsic motivation.

Other than motivation, research has also honed in on instruction as an important variable in L2 pronunciation accuracy. Elliot (1995b) demonstrated that an experimental group receiving phonetics lessons produced Spanish phones significantly more accurately than a matched control group after a semester of instruction (see also Lord, 2005). Moyer (1999) found that a combination of segmental and suprasegmental training benefited phonological accuracy in L2 German. Other studies, however, evidence a more complex relationship between accuracy and training. Helmke and So (1980) found that aptitude interacted with training type; high aptitude learners appeared to benefit more from pronunciation exercises compared to low aptitude learners who benefited more from single drills. Shively (2008) found that class level and amount of formal instruction in Spanish were significantly related to participants' production of Spanish [β], [ð], and [ɣ]. Two groups of participants recruited from a second semester Spanish class or a phonetics class recorded two speaking tasks: a set of words and a paragraph targeting the intervocalic voiced approximants [β], [ð], and [ɣ]. The phonetics students were more accurate overall than the second semester students, and participants with more formal instruction in Spanish outperformed participants with less instruction when the participant groups were analyzed together but not when each group was considered individually.

Aside from motivation and instruction, researchers have investigated the role of language immersion (Díaz-Campos, 2004; Moyer, 1999; Shively, 2008; Simões, 1996) and phonological short-term memory (PSTM) (Aliaga-García et al., 2011; MacKay et al., 2001) in L2 phonological accuracy. There is a wealth of research on language immersion for naturalistic learners—adults learning a foreign language through direct exposure to the language, typically living in a country where the target language is spoken. For naturalistic learners, research appears to indicate that length of residence in the country is a poor predictor of foreign accent compared to learners' age of arrival (see, e.g. Flege, 1999; Trofimovich & Baker, 2006). Research on adults learning an L2 in an instructional setting also presents mixed results regarding language immersion. Moyer (1999) found that years of language immersion did not correlate with better accuracy for adult English speakers of L2 German. Shively (2008) obtained similar findings: Differences in length of immersion for the phonetics group did not reach significance. Learners who had studied abroad for longer actually performed worse; in fact, the learner who had studied abroad for five months was the least accurate. Díaz-Campos (2004) also failed to find positive results for a language immersion group. He compared a study-abroad group to a control group after a semester spent abroad; the study-abroad group performed similarly to the control. Despite these bleak results, it is important to bear in mind that the characteristics of an immersion program vary considerably both between programs and from participant to participant. Without detailed information on the study-abroad program itself (e.g. L2 use, amount of interaction with native speakers, etc.), researchers cannot attest to the authenticity of the immersion experience. It may be that a true immersion program where learners are engaged in meaningful conversations with native speakers on a daily basis does help improve pronunciation. One study that appears to support that conclusion is Simões (1996) who found that learners of L2 Spanish made improvements in vowel quality after a five week study-abroad program in Costa Rica.

Phonological short-term memory, part of Baddeley and Hitch's model of working memory (1974), refers to a temporary storage system for oral input; crucially, in order for information to remain in PSTM, it must be continuously refreshed through subvocal rehearsal (Baddeley, 2003). Like the other factors outlined above, previous research has shown PSTM to be important for L2 phonology, especially for L2 speech perception (Aliaga-García et al., 2011; Cerviño-Povedano & Mora, 2010; MacKay et al., 2001). In MacKay, Meador, and Flege (2001), Italian-speaking learners of L2 English identified word-initial and word-final consonants in spoken disyllables by selecting the appropriate response from five written options presented in standard English orthography. The learners also took a nonword repetition test in Italian as a measure of their PSTM. Results revealed that PSTM was negatively correlated with error scores on the dependent measures ($r = -0.29$ word-initial consonants, $r = -0.42$ word-final consonants); learners with higher PSTM made less errors when identifying both initial and final consonants. PSTM scores also accounted for 15 and 8% of the variance in identification error scores for word-final and word-initial consonants, respectively. The authors

suggest that PSTM may be related to the ability to establish phonetic categories for L2 sounds in that learners with higher PSTM are perhaps more able to create long-term representations of L2 consonants. Similarly, Cerviño-Povedano and Mora (2010) and Aliaga-García, Mora, and Cerviño-Povedano (2011) both explored the relationship between PSTM and cue weighting in L2 English; that is, they investigated the extent to which Catalan-speaking learners of L2 English relied on duration as an indicator of English vowel distinctions such as /i:/–/i/. Previous research has shown that whereas native English speakers rely on both duration and spectral quality to distinguish vowels, nonnative speakers tend to privilege durational cues. Using median splits of PSTM scores on serial nonword recognition tasks, the researchers divided learners into high and low PSTM groups. Learners with higher PSTM did not rely on duration as much as learners with low PSTM (Cerviño-Povedano & Mora, 2010; Aliaga-García, Mora, & Cerviño-Povedano, 2011). Furthermore, low PSTM learners were especially reliant on duration for synthesized vowels—vowels whose duration had been manipulated along a continuum of equidistant steps (Aliaga-García, et al., 2011). This research follows the hypotheses put forth in MacKay, Meador, and Flege (2001): Learners with higher PSTM may create more accurate long-term representations that better approximate native perception than learners with lower PSTM.

To generalize across the diverse methodologies and populations under consideration in these studies is to arrive at a tentative and incomplete view of the factors implicated in L2 phonological attainment: Motivation is important but the exact nature of that motivation remains unclear, instruction appears to benefit pronunciation but may be confounded with amount of language experience, language immersion may help learners to improve their pronunciation depending on the characteristics of the program, and higher PSTM appears to relate to the ability to correctly perceive L2 phones. The current study seeks to contribute to this picture by investigating what variables best predict degree of foreign accent in L2 Spanish. The goal of this research is to further our understanding of the specific nature of the variables implicated in high-proficiency or near-native L2 speech production.

1.2. Is there a critical period for L2 phonology?

Related to research on variables implicated in L2 phonological development is the notion of a critical period for language acquisition; can a learner attain a near-native or native-like accent in a foreign language after a given window of opportunity? Critical period research clearly demonstrates that the probability of a near-native L2 phonology rapidly diminishes as we age (Flege, 1987a; Patkowski, 1980; 1990; Scovel, 1988). Some studies even report that children acquiring an L2 before the age of 6 have a discernible foreign accent (Flege, 1999; Oyama, 1976). Abrahamsson and Hyltenstam (2009) conducted one of the most rigorous studies to date on the issue of ultimate attainment in a foreign language and came to the conclusion that native speakers of Spanish who were late-learners of L2 Swedish did not perform like native speakers on a series of complex linguistic tasks. Even learners with a low age of arrival were unable to match the native-speaker control subjects. Yet the researchers required potential participants to interview with native Swedish speakers and pass as natives before participating in the experiment; the learners in fact were capable of passing as native speakers if passing is defined through native-speaker perception and not linguistic tasks. Moyer (1999) obtained similar results in her study. Participants who were late-learners of L2 German completed speaking tasks in the L2 which native judges rated. Judges rated one of the late-learners as a native speaker when in fact this learner had spent little time abroad and had studied German for a relatively short interval compared to the other subjects. Therefore, it seems that native-like achievement is possible for learners in both naturalistic (Abrahamsson & Hyltenstam, 2009) and instructional (Moyer, 1999) contexts despite the early onset of age effects in L2 phonology. A second goal of this study then is to attempt to identify near-native learners—learners who can pass as a native speaker among native speakers of the L2. These learners are critical data points for the factors that most influence phonological ultimate attainment as they should possess some characteristics in common that allow researchers to identify what really matters for phonological development. Consequently, this study draws data from two independent sources: the factors that predict L2 phonological accuracy for the overall participant pool and exceptional learners as a case study for what truly fosters greater pronunciation accuracy.

2. Method

2.1. Participants

Participants were adult learners ($n = 34$, $M_{age} = 21.94$, age range: 18 – 33 years) who had taken at least six semesters of college-level Spanish or the equivalent thereof and adult English-Spanish bilinguals ($n = 3$). Of an initial pool of 42 participants, five learners of Spanish were excluded for the following reasons: (a) English was not their first language; (b) prior knowledge of the nonwords on the PSTM task (i.e. knowledge of Russian); (c) failure to complete at least one of the assessment tasks.

2.2. Raters

The researcher initially recruited eight native speakers of various Spanish dialects as raters. None of the raters reported any formal linguistic or phonetic training on a background questionnaire administered prior to the rating session. Previous research has shown that naïve, i.e. untrained, and expert raters do not significantly differ in the ratings given (Bongaerts, van Summeren, Planken, & Schils, 1997). Upon reviewing the data and in personal interviews with the raters after the experiment, the researcher decided to exclude ratings from four raters because they failed to adhere to the rating procedure or misinterpreted the scale.

2.3. Materials

2.3.1. Language background questionnaire

The background questionnaire elicited language contact data (Appendix 7.1). Participants reported on the amount and type of Spanish instruction and other foreign languages known (see Table 1).

Table 1

Descriptive Statistics for Learners of L2 Spanish

| Variable | <i>M</i> | <i>SD</i> | Range |
|---------------------------------------|----------|-----------|----------|
| Age | 21.94 | 3.88 | 18 – 33 |
| Age of Exposure | 10.76 | 4.63 | 3 – 24 |
| Age of Formal Study | 13.21 | 3.09 | 7 – 24 |
| Years of Formal Study | 8.46 | 2.81 | 3.5 – 15 |
| Months of Immersion | 5.11 | 6.24 | 0 – 26 |
| Years of Spanish Teaching Experience | 0.74 | 1.60 | 0 – 6 |
| Outside Spanish Immersion (hrs./week) | 4.13 | 9.33 | 0 – 40 |
| Other Foreign Languages Known | 1.62 | 1.63 | 0 – 6 |

2.3.2. Motivation assessment

In this study, motivation was operationalized following the Motivational Self System (Dörnyei, 2010). Motivation in this model has three components: the *ideal L2 self*, the *ought-to L2 self*, and the *L2 learning experience*. The *ideal L2 self* responds to desire; learners envision a future ideal self that is a composite of what they seek to attain in a foreign language and attempt to reconcile their current competency with that future self. The *ought-to L2 self* is naturally born out of obligation; what is it that a learner feels he should accomplish in a foreign language for personal or professional purposes?

In this study, only the *ideal* and *ought-to L2 self* components of the model were measured. Participants completed a 16-item survey assessing both L2 selves; ten questions assessed the *ideal L2 self* and six the *ought-to L2 self*. For each item, participants indicated how much they agreed or disagreed with a statement related to language learning using a Likert-type, six-point scale (1 = *strongly disagree*, 6 = *strongly agree*) (see Appendix 7.2). The researcher based this assessment on a similar questionnaire provided in Dörnyei (2010).

2.3.3. *Phonological short-term memory task*

Participants took a nonword pairs repetition test (NWR) in Russian (Stafford, 2005) to assess their PSTM. The test consisted of 16 nonword pairs at two, three, four, and five syllable lengths. All nonwords conformed to the phonotactics of Spanish and English; tautosyllabic consonant clusters legal in Russian (e.g. [stv]) but illegal in English or Spanish were avoided. A female native speaker of American English recorded the words for the NWR task. Participants received instructions informing them that they would hear a pair of words and repeat them as accurately as possible after a tone. Participants had approximately six seconds to repeat the words after the tone before the program advanced to the next pair. Recordings were made in QuickTime and uploaded to a protected site where the researcher could later review them.

2.3.4. *Operation-span task*

The researcher evaluated participants' working memory (WM) capacity using an operation span (OSPAN) task (Medina, 2008). Participants completed a WM measure because greater working memory may allow learners to process, store, and recall oral input in greater scrutiny. The OSPAN task consisted of a four set training phase and an 18 block test phase. Participants solved a time-controlled series of simple arithmetic problems and attempted to recall the addends of the series in the order in which they appeared at the end of each block upon seeing RECALL. The number of items per block (1–9) increased throughout the task. The OSPAN task was automatized using PowerPoint and progressed automatically; participants had a set amount of time to solve each arithmetic problem before the slideshow advanced.

2.3.5. *Speaking tasks*

Participants' scores on the three speaking tasks were the dependent variables of the study. On the word task, participants read aloud a series of ten words containing segments predicted to be challenging for L1 English speakers of L2 Spanish based on both the predictions of current models of L2 phonological learning such as Flege's Speech Learning Model (1992, 1995) or Best's Perceptual Assimilation Model (1995) and a review of teaching materials designed for learners of L2 Spanish (Guitart, 2004; Hualde, 2005). The words targeted the following phones: voiced approximants, syllable-initial voiceless stops, rhotics, and diphthongs. On the sentence task, participants read aloud a paragraph at a comfortable pace. Lastly, participants responded to a prompt in Spanish on the free response task. These tasks generally followed Moyer (1999) (Appendix 7.3).

2.3.6. *Rating tasks*

The native Spanish-speaking raters assessed each sound file on an eight-point scale, the higher the rating, the more nativelike the accent. Southwood & Flege (1999) determined that a nine-point scale is ideal for assessing foreign accent. An eight-point scale was ultimately selected to force raters to rate each sample as more or less accented—on an eight-point scale there is no obvious midpoint.

2.4. *Procedure*

Data was collected across two sessions. In the first session, participants completed the language background questionnaire, motivation assessment and PSTM and OSPAN tasks. In a second individual session with the researcher, they recorded the word, paragraph, and free response tasks in that order.

2.4.1. *Speaking tasks*

For the speaking tasks, the researcher met with participants individually. Before each task, participants were instructed to silently read through the text to familiarize themselves with what they would read aloud. This was done to minimize unnatural pausing. For the word task in particular,

participants were instructed to pause briefly after each word. Recordings were made in Audacity using a high quality microphone attached to a laptop computer.

2.4.2. Rating tasks

The researcher edited participants' recordings to prepare them for the rating task by eliminating fluency errors (e.g. unnatural pausing) that could influence the raters' decisions. On the word task, the researcher edited the recordings to include the first full articulation of each item; false starts (e.g. pronouncing only part of a word) were eliminated. Paragraph task recordings were edited exclusively for length; the researcher shortened all recordings to the first three sentences to make the rating task itself shorter and more manageable for the raters. Similarly, the researcher extracted a short excerpt of approximately 30 seconds from participants' free response recordings, avoiding the first and last sentences which typically contained the most disfluent speech.

An E-Prime file was created to randomize the speech samples within each block and present them to the raters. Two optional short-breaks (5 minutes) were built into the file occurring after each block to prevent rater fatigue. Raters were first presented the following instructions in English:

In this study, you are going to listen to speech clips in Spanish that were recorded by both native and nonnative speakers. You are going to rate each clip on a scale of 1 – 8, 1 being a **very strong foreign accent** and 8 being **no foreign accent**.

Please base your ratings solely on pronunciation. Disregard the fluency and grammar of the speakers.

You will listen to a series of words and paragraphs. You will only be able to listen to each clip once so please listen carefully. After listening to each clip, press the space bar to continue to the ratings screen. Do not press the space bar until the clip has finished playing.

You will have two breaks in between tasks. You should expect to spend between an hour and an hour and a half completing the ratings. If you have questions, please consult the researcher now.

Raters pressed one key to begin the task and another to advance to a rating screen after the presentation of each sound file. The program automatically advanced to the next clip after the rater had entered a score. Two of the raters proceeded from the more controlled word and paragraph tasks to the more extemporaneous free response task, and two rated in the opposite direction from extemporaneous to controlled. In general, most raters completed the rating task within an hour.

2.4.3. OSPAN and PSTM scoring

Each nonword on the NWR test was worth one point (maximum score = 32). The researcher awarded credit if a participant's repetition of the item did not differ by more than one phoneme from the original recording and if it was correctly stressed—participants could make one phonemic error but stress errors were fatal. The researcher chose to penalize learners for incorrect lexical stress in order to ensure that they paid sufficient attention to the suprasegmental characteristics of the target nonwords.

OSPAN scoring followed recommendations outlined by Conway, Kane, Bunting, Hambrick, Wilhelm, and Engle (2005): Participants had to meet an 85% accuracy threshold on the processing component (the arithmetic problems) in order for their scores on the recall portion of the task to be considered valid. The scoring of the recall items then adhered to partial-credit unit scoring.

3. Results

First, interrater reliability was assessed by calculating Kappa for each pair of raters and averaging the four Kappa scores to create a composite interrater reliability score for each task. Interrater reliability was fair for the Word, $\kappa = .265$, Paragraph, $\kappa = .229$, and Free Response tasks, $\kappa = .211$, following the Landis and Koch (1977) scale: 0.21 – 0.40 = fair agreement. These scores were then averaged to establish a reliability coefficient for the overall Pronunciation task, $\kappa = .235$. Although these scores appear to be on the low side, it is important to bear in mind that reliability tends to

decrease as categories increase. Considering that this study employed a scale with eight possible categories (1–8), lower interrater reliability is perhaps more acceptable.

Given that the goal of this study was to identify factors that predict overall pronunciation accuracy as opposed to specific aspects of pronunciation (e.g. segmental faithfulness on the word task), the researcher then calculated a composite pronunciation score for each participant. This score was an average of the scores obtained on the three tasks, which were all significantly ($p < .001$) correlated with one another: Word with paragraph, $r = .93$, word with free response, $r = .87$, paragraph with free response, $r = .92$. Descriptive statistics of the overall sample on the speaking tasks are presented in Table 2.

Table 2

Descriptive Statistics of the Overall Sample on the Speaking Tasks

| Task | <i>M</i> | <i>SD</i> | Range |
|---------------------|----------|-----------|-------------|
| Word | 4.01 | 1.96 | 1.00 – 8.00 |
| Paragraph | 3.80 | 1.96 | 1.00 – 7.75 |
| Free Response | 3.69 | 2.14 | 1.00 – 7.75 |
| Overall Mean Rating | 3.84 | 1.94 | 1.25 – 7.83 |

The researcher subsequently conducted two correlational analyses: correlations among independent variables (Table 3) and correlations between independent variables and mean pronunciation rating (Table 4).

Table 3

Significant Correlations between Independent Variables

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---------|---|--------------------|--------------------|---|-------------------|------------------|--------------------|-------------------|-------------------|------------------|-------------------|
| 1 AoE | — | .63 ^{***} | — | — | .51 ^{**} | .43 [*] | .42 [*] | .43 [*] | — | — | — |
| 2 AFS | — | — | -.45 ^{**} | — | — | — | — | — | — | — | — |
| 3 YFS | — | — | — | — | — | .40 [*] | — | — | .39 [*] | — | — |
| 4 MoI | — | — | — | — | — | .38 [*] | .60 ^{***} | — | .44 ^{**} | — | — |
| 5 OFL | — | — | — | — | — | .35 [*] | .51 ^{**} | .50 ^{**} | — | .41 [*] | — |
| 6 YTE | — | — | — | — | — | — | .64 ^{***} | .46 ^{**} | .45 ^{**} | .36 [*] | — |
| 7 OLE | — | — | — | — | — | — | — | .40 [*] | .45 ^{**} | — | — |
| 8 Ideal | — | — | — | — | — | — | — | — | — | — | — |
| 9 Oug. | — | — | — | — | — | — | — | — | — | .39 [*] | -.35 [*] |
| 10 PM | — | — | — | — | — | — | — | — | — | — | — |
| 11 OS | — | — | — | — | — | — | — | — | — | — | — |

AoE = age of (initial) exposure; AFS = age of onset of formal study; YFS = years of formal study; MoI = months of immersion; OFL = number of other foreign languages known; YTE = years of experience teaching Spanish; OLE = outside language exposure, i.e. L2 use outside of an instructional setting; Ideal = ideal L2 self; Oug. = ought-to L2 self; PM = phonological short-term memory; OS = operation span task. * $p < .05$; ** $p < .01$; *** $p < .001$

Table 4

Significant Correlations between Independent Variables and Mean Pronunciation Rating

| Measure | Correlated with | <i>r</i> | <i>p</i> |
|---------------------|-----------------|----------|----------|
| Age of Exposure | Mean Rating | .47 | .005 |
| Ideal L2 self | | .59 | < .001 |
| Ought-to L2 self | | .42 | .013 |
| Months of Immersion | | .65 | < .001 |
| PSTM | | .51 | .002 |

Based on the correlations in Table 4, the researcher conducted a regression analysis on participants' mean pronunciation ratings to assess which factors best predict pronunciation accuracy in L2 Spanish. The analysis revealed three significant predictors: length of immersion ($p = .001$), ideal L2 self ($p = .003$), and phonological short-term memory ($p = .031$) which together explained 70% of the overall variance ($r^2 = .70, p < .001$). Ought-to L2 self and age of exposure were not significant with p values of .759 and .388 respectively. A subsequent hierarchical regression identified length of immersion as the single best predictor of mean rating, followed by ideal L2-self and PSTM (Table 5).

Table 5
Hierarchical Regression on Mean Pronunciation Rating

| Step | Variable | p | r^2 | Δr^2 | Significance of change |
|------|---------------------|--------|-------|--------------|------------------------|
| 1 | Length of Immersion | < .001 | .42 | .42 | < .001 |
| 2 | Ideal L2-self | .001 | .62 | .20 | < .001 |
| 3 | PSTM | .009 | .70 | .8 | .009 |

After analyzing the group data, the researcher then considered the data of individual participants to determine if any L2 learners had achieved a native-like rating (native range = 7.08 – 7.92). Evaluating those participants who had obtained high pronunciation scores was seen as crucial to the goals of this study as it would not only help to clarify how maturational constraints operate in an instructional, i.e. formal learning, context but also provide further insight into the factors identified in the regression analysis. Three learners attained mean ratings that fell within the native range with scores of 7.42, 7.75, and 7.83. These three exceptional learners had acquired Spanish well after the closure of any window of native acquisition and excelled in the factors identified as predictors of mean rating (Table 6).

Table 6
Characteristics of Exceptional Learners

| Learner | Age of Onset | Immersion | Ideal L2 self | PSTM |
|----------|--------------|-----------|---------------|-------|
| 1 | 13 | 26.00 | 59 | 26 |
| 2 | 14 | 8.00 | 60 | 30 |
| 3 | 24 | 21.00 | 58 | 29 |
| <i>M</i> | 17 | 18.33 | 59 | 28.33 |

4. Discussion

4.1. Predictor variables

The primary goal of this study was to identify the factors that predict ultimate attainment in L2 phonology. The analyses revealed three significant predictors of mean pronunciation rating: length of immersion, ideal L2 self, and PSTM. Each of these factors is discussed in turn.

There appears to be an intuitive link between immersion and language learning, particularly as concerns phonology. Most study abroad experiences provide learners with vast amounts of diverse oral input that may be considered a form of implicit phonological training. Furthermore, immersion requires the learner to produce the target language much more often and in more authentic contexts than in the classroom. Through repeated production exercises, learners may have been able to solidify emerging L2 phonological patterns although, to the author's knowledge, no current research has specifically addressed this claim.

What can be said is that, while not a panacea for all phonological problems, the reviewed literature on language immersion does suggest that learners enjoy an initial period of rapid phonological development followed by stabilization (Flege, 1988). In other words, it seems that immersion does help learners acquire L2 phonological patterns, although not necessarily to nativelike levels. Future research should explore the connection between implicit phonological training and

pronunciation accuracy, attempting to tease apart what makes language immersion so special for phonological development.

For the set of participants in this study, language learning is no longer required, but optional. They have completed the requisite number of courses stipulated by the university and most have chosen Spanish as a major and opted to study abroad. Therefore, it stands to reason that in an optional language learning setting, the ideal L2 self is a more reliable predictor of pronunciation accuracy than the ought-to L2 self, which might be more important in naturalistic settings. Because the ideal L2 self concerns learners' ability to envision a future successful self and bridge the gap between their current capabilities and their desired competency, perhaps highly successful learners develop more elaborate and detailed language goals in comparison to their less successful peers. Researchers should continue to probe the relationship between motivation and phonological attainment, focusing on the types of language goals learners generate.

In this study, PSTM was operationalized as the ability to correctly encode and maintain in memory the segmental and suprasegmental features of speech; learners had to correctly recall the phonemes and lexical stress of the nonword pairs. By this criterion, learners with better PSTM may be able to more accurately encode the phonetic details of novel oral input, allowing them to more closely approximate native pronunciation. Past research on speech perception supports this claim (MacKay et al., 2001; Aliaga-García et al. 2011; Cerviño-Povedano & Mora, 2011) insofar as learners with higher PSTM appear to be more capable of establishing accurate long-term representations that are more closely aligned with native perception than learners with low PSTM. This in turn helps them to develop new phonetic categories. In speech production, these learners may be more successful at producing more targetlike phonological output precisely because of these enhanced representations.

Summarizing, it seems that pronunciation accuracy in a foreign language depends upon a constellation of factors. For the learners in this study, length of immersion, motivation, and PSTM proved to be crucial. What remains to be seen is how exactly these factors interface to account for the end-state of the L2 phonological system.

4.2. Exceptional learners

The second research question guiding this study sought to identify late-learners who pass as native speakers. Three learners received ratings that fell within the native range, both on the composite pronunciation score and on the scores they attained on each speaking task. Table 5 already indicated that each learner ranked highly with respect to the three predictor variables. The exceptional participants also began learning Spanish well beyond what would be considered the window of native acquisition at ages 13, 14 and 24.

All three learners were graduate students whose ultimate goal was to attain high proficiency in order to teach at the university level. They also reported that they had competency in multiple foreign languages beyond Spanish and that they consistently spoke Spanish from 15–40 hours per week ($M = 28.3$) with people with whom they had a close relationship. On the other hand, none of them indicated that they had received phonetic training or feedback on their pronunciation.

Although L2 language use—the amount of time participants spent speaking Spanish outside of an instructional setting—was not significantly correlated with mean rating, it appears to be a unifying trait among the exceptional learners. Previous research on L1 and L2 language use has focused on learners in immersion contexts and obtained mixed results. Purcell and Suter (1980) found that a composite variable that grouped years in an English speaking country with months of residence with a native speaker of English was the third best predictor of participants' pronunciation scores in an analysis of late-learners of L2 English. Flege and Fletcher (1992), however, studied a group of native speakers of Spanish who were learners of L2 English and found that participants' percentage daily L2 use was not significantly related to their foreign accent. Despite these mixed results, increased L2 use in authentic contexts with native speakers may be more important for learners in instructional settings who receive far less L2 input and meaningful practice than similar learners in an immersion context. Perhaps L2 use is more important for classroom learners because their exposure to the target language is more limited.

These learners had also not engaged in any pronunciation training. The fact that they had achieved a near-native accent in their L2 without training suggests that, although beneficial, pronunciation instruction is not necessary for nativelike ultimate attainment, at least for some learners. A productive line of inquiry for future research should investigate what type of learner benefits the most from formal instruction.

5. Limitations and future research

Several limitations to this study must be mentioned. Participants carried out the PSTM as a group in a laboratory setting and therefore may have been exposed to some of the nonwords as their cohorts completed the task. Future studies should make every attempt to have participants complete the PSTM task individually. Additionally, future studies may wish to adopt a different PSTM assessment. Normally PSTM tasks are administered in the participants' L1. Some research suggests that language familiarity may influence the outcome of these measures (see, e.g., French & O'Brien, 2008) while others find PSTM to be largely language-independent (Cerviño-Povedano & Mora, 2011). Another consideration is the participant pool. This study and others like it (Moyer, 1999) targeted high proficiency learners. To increase generalizability, future studies should also assess lower proficiency learners. As regards the ratings, factors other than pronunciation (e.g. fluency or grammatical competency) may have influenced the raters although the instructions for carrying out the ratings specifically told them to rate solely based on pronunciation. Future research on pronunciation should make every effort to provide recordings free from other sources of error (e.g., Bongaerts et al., 1997). Finally, the data on formal pronunciation instruction which was collected on the language background questionnaire was difficult to address given that participants frequently gave vague or incomplete responses. Future research must take great care in designing questionnaires or other materials that collect similar data.

6. Conclusion

This study was designed to investigate the factors that predict ultimate attainment in second language phonology. Participants completed a number of tasks which evaluated their amount of language experience, type and level of motivation, and cognitive abilities. They then recorded a series of controlled and extemporaneous speaking tasks designed to capture their natural speech patterns. Native Spanish-speaking raters subsequently listened to the speech files and assigned a rating on an eight-point scale. Correlational and regression analyses revealed three significant predictors of mean pronunciation rating: length of immersion, ideal L2 self motivation, and PSTM. A more detailed analysis of the exceptional learners revealed that, apart from excelling in the factors identified in the regression analyses, they also shared similar language learning goals and frequently used the L2 for personal reasons. The results of this study (a) highlight the fact that learners appear to be able to achieve near-native pronunciation without significant formal instruction in pronunciation which perhaps also evidences the role of implicit learning in L2 phonology; (b) establish some of the factors that seem to account for the successful language learner as far as phonology is concerned; (c) provide a basis for future empirical research probing the relationship between pronunciation accuracy and length of immersion, motivation, and PSTM.

7. Appendices

7.1. Background questionnaire

Please take your time answering each question, as your honesty and insights for each question are invaluable for drawing conclusions from this project. Thank you!

Part One: Background

1. Age
2. Sex
3. Do you have any known visual or hearing impairments?
4. What is your native language (the language you grew up speaking)?
5. Did you speak any other languages at home as a child? If YES, please list them.
6. Have you studied any other foreign languages, excluding Spanish?
7. Please list any foreign language, excluding Spanish, that you have any knowledge of.
8. Please briefly describe the nature of your experience with each language you mentioned.

Part Two: Spanish Instruction

1. Indicate the age at which you were first exposed to Spanish.
2. Indicate the age at which you first formally studied Spanish.
3. Indicate the years of total instruction in Spanish you have had up until now.
4. Have you ever visited or studied in a Spanish-speaking country?
5. For each experience, please list (1) the country (2) nature and (3) length of your stay.
6. Indicate the total amount of time you have spent in an immersion environment.
7. Have you ever taught Spanish?
8. If YES, indicate the total number of years you have spent teaching Spanish.
9. Does anyone with whom you have a close relationship speak Spanish?
10. Do you consistently speak with that person or those people in Spanish?
11. List the approximate hours per week you spend speaking Spanish with those people.
12. List your primary motivation for studying Spanish, either professional or personal.
13. List your ultimate goal in studying Spanish.

Part Three: Phonological Profile

1. Have you taken a course dedicated to improving your Spanish pronunciation?
If YES, please list the course(s), the type(s) of instruction (drills, contrasting English and Spanish, etc.) and the focus of instruction (specific sounds, the rhythm of Spanish, etc.). Please be as detailed as possible.
2. Do you feel that you improved as a result of the course? Please explain your response.
3. Have you practiced pronunciation skills in any other course you have taken?
If YES, please list the course(s), approximately how much time (in hours) was dedicated to pronunciation on a weekly basis, the type(s) of instruction and the focus of instruction.
4. Do you feel that you improved as a result of that instruction? Please explain your response.
5. Have you ever gotten any feedback on your pronunciation outside of class?
If YES, please list who gave you the feedback (native vs. nonnative Spanish speaker), whether or not you consistently received feedback, and the focus of the feedback (one particular sound, a word, general rhythm of the language).
6. Do you feel that the feedback was helpful or that you improved as a result? Please explain your response.
7. Overall, has the instruction and/or feedback you received focused more on specific sounds or the rhythm and flow of Spanish?
8. Overall, do you feel that the instruction and/or feedback has been helpful? Why or why not?

7.2. Motivation assessment

In this survey, I would like you to tell me how much you agree or disagree with the following statements by simply circling a number from 1 to 6. Please do not leave out any items.

| | Strongly disagree | Disagree | Slightly disagree | Slightly agree | Agree | Strongly agree |
|-----|-------------------|----------|-------------------|----------------|-------|----------------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1. | | | | | | |
| 2. | | | | | | |
| 3. | | | | | | |
| 4. | | | | | | |
| 5. | | | | | | |
| 6. | | | | | | |
| 7. | | | | | | |
| 8. | | | | | | |
| 9. | | | | | | |
| 10. | | | | | | |
| 11. | | | | | | |
| 12. | | | | | | |
| 13. | | | | | | |
| 14. | | | | | | |
| 15. | | | | | | |
| 16. | | | | | | |

7.3. Speaking tasks

You will have a series of four tasks, and you will pause between each task. You will complete the tasks in the order presented, and the researcher will record your responses. You will be able to record your responses only once as re-recording is a form of practice.

Please read the following words at a natural pace:

- | | |
|---------------|--------------|
| 1. Monstruo | 6. Tarjeta |
| 2. Guerra | 7. Aduana |
| 3. Vaso | 8. Funcionar |
| 4. Pintura | 9. Examen |
| 5. Partidario | 10. Vidrio |
-

Please read the following passage at a comfortable pace.

Las legumbres han sido uno de los principales sustentos de la población en épocas de escasez. Se debe a su fácil conservación, su alto valor nutritivo y su gran versatilidad en la cocina. Hasta hace poco, constituían la principal comida del día, habitualmente acompañadas de embutidos y carne, excepto los viernes de vigilia, en que se acompañaban de verduras y bacalao. Los vegetales son de fácil cultivo ya que son poco exigentes en cuanto al clima y al suelo, y mejoran la fertilidad del terreno. Sin embargo, a pesar del interesante valor nutricional de las legumbres, en los últimos años se observa una disminución del consumo de vegetales.

Please respond to ONE of the following “guided communication” items in Spanish (or substitute one of your own). Your response need be only 5–10 sentences.

- A. Describe your weekend or your daily routine: What you normally do, when, with whom, for how long, what’s interesting about it, etc.
- B. Describe an experience you had which was meaningful in your life: Who was involved? How old were you? How did this influence you?
- C. Describe a person in your life who means a lot to you: How do you know this person? Why is he/she significant in your life?
- D. Describe a problem or challenge you recently faced and how you dealt with it: What steps did you take to solve it? What was the outcome? Who was affected?
- E. Discuss an issue or subject matter you are interested in: Why is this important for your life? How did you become so interested in it? What has shaped your views of it?

References

- Aliaga-García, Cristina, Mora, Joan C., & Cerviño-Povedano, Eva (2011). L2 speech learning in adulthood and phonological short-term memory. *Poznán Studies in Contemporary Linguistics*, 47(1), 1–14.
- Abrahamsson, Niclas, & Hyltenstam, Kenneth (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, 30, 482–509.
- Abrahamsson, Niclas, & Hyltenstam, Kenneth (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language learning*, 59(2), 249–306.
- Baddeley, Alan (2003a). Working memory and language: An overview. *Journal of Communication Disorders*, 36(3), 189–208.
- Baddeley Alan (2003b). Working memory: looking back and looking forward. *Nature Reviews. Neuroscience*, 4(10), 829–839.
- Baddeley, Alan, & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *Psychology of learning and motivation*, Vol. VIII (pp. 47–90). New York: Academic press.
- Best, Catherine T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore: York Press.
- Bongaerts, Theo, van Summeren, Chantal, Planken, Brigitte, & Schils, Erik (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447–465.
- Cerviño-Povedano, Eva, & Mora, Joan C. (2011). Investigating Catalan-Spanish bilingual EFL learners’ over-reliance on duration: Vowel cue weighting and phonological short-term memory. In M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in SLA of speech: New Sounds 2010, Volume 1* (pp. 53–64). New York: Peter Lang.

- Conway, Andrew R. A., Kane, Michael J., Bunting, Michael F., Hambrick, D. Zach, Wilhelm, Oliver, & Engle, Randall W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769–786.
- Dörnyei, Zoltan (2010). *Questionnaires in second language research: Construction, administration and processing*. New York: Routledge.
- Elliot, A. Raymond (1995a). Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *Modern Language Journal*, 79(3), 356–371.
- Elliot, A. Raymond (1995b). Foreign language phonology: field independence, attitude, and the success of formal instruction in Spanish pronunciation. *Modern Language Journal*, 79(3), 530–542.
- Flege, James E. (1987a). A critical period for learning to pronounce foreign language. *Applied Linguistics*, 8, 162–177.
- Flege, James E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustic Society of America*, 84(1), 70–79.
- Flege, James E. (1992). Speech learning in a second language. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 565–604). Baltimore: York Press.
- Flege, James E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–272). Baltimore: York Press.
- Flege, James E., & Fletcher, Kathryn L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91, 370–389.
- Flege, James E., Yeni-Komshian, Grace H., & Liu, Serena (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41, 78–104.
- French, Leif M., & O'Brien, Irena (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, 29, 463–487.
- Guitart, Jorge M. (2004). *Sonido y sentido: Teoría y práctica de la pronunciación del español con audio CD*. Washington: Georgetown University Press.
- Hualde, José Ignacio (2005). *The sounds of Spanish*. Cambridge: Cambridge University Press.
- Landis, J. Richard, & Koch, Gary G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Lord, Gillian (2005). (How) can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, 88(3), 557–567.
- MacKay, Ian R. A., Meador, Diane, & Flege, James Emil (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, 58, 103–125.
- Medina, Almitra (2008). *Concurrent verbalization, task complexity, and working memory: Effects on L2 learning in a computerized task* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3315458)
- Moyer, Alene (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, 21, 81–108.
- Muñoz, Carmen, & Singleton, David (2007). Foreign accent in advanced learners. *EUROSLA Yearbook*, 7, 171–190.
- Oyama, S. C. (1976). A sensitive period for the acquisition of a phonological system. *Journal of Psycholinguistic Research*, 5, 261–283
- Patkowski, Mark (1980b). The sensitive period for the acquisition of syntax in a second language. *Language Learning*, 30, 449–472.
- Patkowski, Mark (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, 11, 73–89.
- Piske, Thorsten, MacKay, Ian R. A., & Flege, James E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, 191–215.
- Purcell, Edward T., & Suter, Richard W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning*, 30, 271–287.
- Scovel, Thomas (1988). *A time to speak. A psycholinguistic inquiry into the critical period for human speech*. Cambridge, MA: Newbury House.
- Simões, Antônio R. (1996). Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish. *Hispania*, 79(1), 87–95.
- Shively, Rachel L. (2008). L2 acquisition of [β], [ð], and [ɣ] in Spanish: Impact of experience, linguistic environment, and learner variables. *Southwest Journal of Linguistics*, 27(2), 79–114.
- Southwood, Helen M., & Flege, James E. (1999). Scaling foreign accent: Direct magnitude estimation versus interval scaling. *Clinical Linguistics & Phonetics*, 13(5), 335–349.
- Stafford, Catherine (2005). *Bilingualism, cognitive capacity and age: A computer-based study in L3 processing* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3230994)
- Trofimovich, Pavel & Baker, Wendy. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1–30.

Selected Proceedings of the 2011 Second Language Research Forum: Converging Theory and Practice

edited by Erik Voss,
Shu-Ju Diana Tai, and Zhi Li

Cascadilla Proceedings Project Somerville, MA 2013

Copyright information

Selected Proceedings of the 2011 Second Language Research Forum:
Converging Theory and Practice
© 2013 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-458-4 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Nagle, Charles. 2013. A Reexamination of Ultimate Attainment in L2 Phonology: Length of Immersion, Motivation, and Phonological Short-Term Memory. In *Selected Proceedings of the 2011 Second Language Research Forum*, ed. Erik Voss et al., 148-161. Somerville, MA: Cascadilla Proceedings Project.
www.lingref.com, document #2913.