SSLA, **31**, 533–557.
doi:10.1017/S0272263109990015

THE RELATIONSHIP BETWEEN L1 FLUENCY AND L2 FLUENCY DEVELOPMENT

Tracey M. Derwing
University of Alberta

Murray J. Munro
Simon Fraser University

Ronald I. Thomson
Brock University

Marian J. Rossiter
University of Alberta

A fundamental question in the study of second language (L2) fluency is the extent to which temporal characteristics of speakers' first language (L1) productions predict the same characteristics in the L2. A close relationship between a speaker's L1 and L2 temporal characteristics would suggest that fluency is governed by an underlying trait. This longitudinal investigation compared L1 and L2 English fluency at three times over 2 years in Russian- and Ukrainian- (which we will refer to here as Slavic) and Mandarin-speaking adult immigrants to Canada. Fluency ratings of narratives by trained judges indicated

We first thank the Slavic language and Mandarin participants as well as the staff at NorQuest and Metro Colleges in Edmonton for their ongoing cooperation. Thanks also to Larissa Gutland and Adele Tan, who provided the translations, transcriptions, and coding of the first language data, and to the group of research assistants who coded the second language temporal variables. We are grateful to Bruce Derwing for his insightful comments. This research was funded by three Social Sciences and Humanities Research Council grants awarded to Tracey Derwing, Murray Munro, and Marian Rossiter.

Address correspondence to: Tracey M. Derwing, Department of Educational Psychology, 6-102 Education North, University of Alberta, Edmonton, AB T6G 2G5, Canada; e-mail: tracey.derwing@ualberta.ca.

© 2009 Cambridge University Press 0272-2631/09 $15.00

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

a relationship between the L1 and the L2 in the initial stages of L2 exposure, although this relationship was found to be stronger in the Slavic than in the Mandarin learners. Pauses per second, speech rate, and pruned syllables per second were all related to the listeners' judgments in both languages, although vowel durations were not. Between-group differences may reflect differential exposure to spoken English and a closer relationship between Slavic languages and English than between Mandarin and English. Suggestions for pedagogical interventions and further research are also proposed.

One of the most noticeable dimensions along which second language (L2) learners differ is fluency—that is, temporal aspects of oral production that influence the degree of fluidity in speech (e.g., pauses, hesitation phenomena, speech rate). This interpretation of fluency is distinct from the notion of fluency as overall proficiency (Chambers, 1997; Lennon, 1990; Skehan & Foster, 1999; Trofimovich & Baker, 2006; Wood, 2001). Even native speakers' temporal fluency can differ quite widely, as evidenced by popular references to fast and slow talkers. The problems in defining oral fluency are well known (Riggenbach, 2000). Although this term denotes an automatic procedural skill on the part of the speaker, fluency must also be considered from the standpoint of the listener (Guillot, 1999). For this reason, the study of fluency necessarily entails an examination of both the temporal features of speech and their relationship to listener judgments. The present study focuses on these issues. However, because we are examining relatively low-proficiency learners, we are not concerned with Fillmore's (1979) higher features of native language fluency such as creativity but rather with the basic dimension of talking with "few pauses" (p. 93).

Because fluency is a defining feature of an individual's first language (L1) output, one might expect that it is a trait (i.e., a relatively permanent characteristic specific to an individual) and that a relationship between the temporal characteristics of a speaker's L1 and L2 productions can be attributed to cognitive factors such as working memory (see Towell & Dewaele, 2005). For example, it seems unlikely that a person who is relatively dysfluent in his or her L1 would be highly fluent in a L2 (Raupach, 1980). An alternative view, however, is that fluency is a language-specific state, such that a speaker could be highly fluent in one language but not fluent at all in another, compared with others of the same general proficiency level. Of course, in reality, there may be aspects of fluency that are traitlike and others that are statelike. A clear understanding of L2 fluency requires an examination of all the factors that influence it, for both theoretical and pedagogical reasons. In view of the complex array of such influences on L2 fluency, we

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

cannot expect to resolve the emerging issue of state versus trait in a single study. Nonetheless, we hope to contribute to the discussion about the relationship between L1 and L2 fluency through a longitudinal investigation of two L1 groups of adult immigrant English-as-a-second-language (ESL) students.

MODELS OF FLUENT SPEECH PRODUCTION

Current conceptions of speech production have been strongly influenced by Levelt's (1989) and Levelt, Roelofs, and Meyer's (1999) model, a L1 framework that postulates that speakers have a general knowledge component; a conceptualizer, in which messages are generated; a formulator, in which grammatical and phonological encoding takes place once lemmas and forms have been accessed in the lexicon; an articulator that produces overt speech; and a speech comprehension system. In L1 speakers, production is highly automatized and rapid speech is the norm. This steady-state model has been adapted by de Bot (1992) for bilingual speech production. He noted that a bilingual model needs to account for a variety of phenomena characteristic of L2 speech, including codeswitching, crosslinguistic influences, slower L2 speech production, and differences in L1 and L2 mastery. Additionally, he observed that a complete model should allow for multiple languages and interactions across those languages. More recently, Kormos (2006) developed a comprehensive model of L2 speech production, also based on Levelt's work, which takes into account previous theoretical constructs and empirical research. She argued that L2 speakers' degree of automaticity affects the rate of speech production in the L2. Furthermore, Kormos's model posits an additional "store of declarative knowledge of syntactic and phonological rules" (p. 178). Although these models are helpful in conceptualizing L2 speech production, they have not yet yielded much insight into the state versus trait issue. Moreover, although it has been argued that the site of L2 fluency problems is the formulator, because that is where declarative knowledge is converted to procedural knowledge (Towell, Hawkins, & Bazergui, 1996), the possibility that L2 oral fluency is also affected in the articulator and the conceptualizer cannot be dismissed.

Insight into the state versus trait issue can be gained through empirical investigation of automatic processing in receptive abilities in both the L1 and the L2 (Segalowitz & Hulstijn, 2005) as well as through a comparison of fluency measures (e.g., acoustic measures and rating data) of L1 and L2 productions. In general, a state view suggests that L1 and L2 fluency are relatively independent, whereas a trait view entails that L1 and L2 fluency are closely related. Although the data discussed here can be used to explore the relationship between L1 and L2 fluency, because

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

we are only concerned with articulatory productions, we will not address where in the speech production process fluency problems arise.

L2 FLUENCY RESEARCH

SLA fluency research has primarily focused on the development of L2 fluency. Most studies have examined only one L1 group at a time—sometimes with very few learners (e.g., Lennon, 1990), often only from a foreign language learning perspective (e.g., Towell, 2002), and often when the L1 and the L2 are related (e.g., English and another Indo-European language). For instance, Segalowitz and Freed (2004) studied English-speaking learners of Spanish. A difference between the current study and others (e.g., Díaz-Campos, 2004; Freed, Segalowitz, & Dewey, 2004; Lafford, 2004) is the availability of detailed longitudinal data from more than two points in time that include both temporal measures of speech and listener evaluations. Although Trofimovich and Baker (2006) studied fluency measures of Korean speakers of English at 3 months, 3 years, and 10 years after arrival in the United States, their study was cross-sectional rather than longitudinal and did not include fluency ratings of L1 productions.

A few previous studies have directly compared L1 and L2 fluency; however, because of methodological differences, their relevance to the present work is limited. Riazantseva (2001) focused on the relationship between L2 proficiency and pausing in both the L1 (Russian) and the L2 (English). The intermediate learners in her study showed more L1-like pause durations in English than did the advanced learners. She concluded that some differences in L1 and L2 fluency are due to a transfer of L1 pausing patterns that "could be language specific" (p. 522) but that increased proficiency in the L2 may allow speakers to overcome the transfer effects. Again, Riazantseva's study differs from the present one in that she examined learners who were of intermediate proficiency or higher, whereas our participants were relative beginners at the outset of the study. Additionally, Riazantseva's study focused entirely on pause phenomena, with no global assessments of fluency. Finally, her pause measures started at 100 ms, a much shorter criterion than the one used here.

Towell and Dewaele (2005) provided a detailed description of conceptual frameworks that have been invoked to account for oral fluency development in a L2. Relevant models make reference to such constructs as automatization (Segalowitz, Segalowitz, & Wood, 1998), differences in working memory and ability to store knowledge (Just & Carpenter, 1992; Pienemann, 1998), neurobiological factors (Dewaele, 2002), and the distinction between declarative and procedural knowledge (Ullman, 2001). In a study designed to evaluate the role of these

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

factors in L2 fluency, Towell and Dewaele observed a positive correlation between L1 and L2 speaking rates in adult learners of French who studied abroad; however, they found slightly higher Spearman correlations during the first or second year of study than after study abroad. Although L2 fluency improved over time in these learners, Towell and Dewaele's hypothesis that differences in working memory would account for the results was not confirmed. Nonetheless, students whose speaking rates remained slow after study abroad performed less well on a recall task in their L1, which suggests that limited ability to store knowledge may have been a contributing factor. Another key finding of Towell and Dewaele's investigation was that improvement in grammatical knowledge, which occurred for all learners, was not necessarily linked to improvement in oral fluency.

Segalowitz and Freed (2004) examined both oral proficiency (ratings on the ACTFL Oral Proficiency Interview) and fluency (i.e., fluidity or smoothness) in learners of Spanish under two conditions: at home in an English-speaking environment and in Spain. Although the study-abroad students made superior progress in fluency development and oral proficiency, the relationship to the learning context was not straightforward. Not only did language contact time fail to predict gains in performance but a negative relationship between hours spent with homestay families and oral proficiency was found. The authors speculate that repetitive, somewhat superficial interactions in the homestay may have led some students to develop compensatory strategies rather than improved linguistic skills. Additionally, Segalowitz and Freed found that "initial oral performance levels may also influence learners' predispositions to make use of extracurricular communicative activities" (p. 195); that is, higher proficiency learners were more likely to engage in extracurricular activities.

Study-abroad research provides important yet often difficult-to-interpret information about the role of context in language learning (Collentine & Freed, 2004). However, the generalizability of this type of research is limited because it centers on participants who have self-selected to study a new language. Immigrants, the focus of the present study, represent a different, more heterogeneous category of learners who are also immersed in a L2 context. They have not self-selected for language study, but rather most have chosen to move to a new country for reasons of employment, family, or security. Because of this difference, it is important not to overgeneralize any aspects of L2 learning, including fluency, on the basis of study-abroad research alone.

RATIONALE FOR THE CURRENT STUDY

This study is part of a series of investigations that examine the development of oral production and perception in Russian- and Ukrainian- (which

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

we will refer to here as Slavic) and Mandarin-speaking adult L2 learners of English. Previous research has identified within-learner cognitive influences on L2 fluency, such as phonological memory (O'Brien, Segalowitz, Freed, & Collentine, 2007), attention (Ellis, 2006), and facility in lexical access (Segalowitz, 2007; Segalowitz et al., 2004). In general, even proficient and fluent L2 speakers tend to speak more slowly than native speakers (Munro & Derwing, 2001); however, relatively little is known about the development of fluency over time. The current study examines the relationship between L1 and L2 fluency productions in light of influences such as overall exposure to spoken English, which is both a precursor and a prerequisite to oral fluency development. Although this influence is external to the learner, to some degree it is within the learner's control. Indeed, Segalowitz observed that "it is reasonable to assume that a person normally develops high levels of access fluidity and attention control only through extensive exposure to and practice with the target language in naturalistic communicative situations" (p. 184).

In our previous work, observations and tests were conducted at approximately 2-month intervals throughout the participants' first year of residence in Canada and at an additional point about 2 years after the outset of the study. In the first investigation, we (Derwing, Rossiter, Munro, & Thomson, 2004) collected fluency ratings from phonetically untrained listeners of the Mandarin learners' productions at 2 months into the study on three distinct tasks. We also measured several temporal factors typically used in research of this type (Towell et al., 1996), such as pauses, mean length of run (i.e., the number of words between pauses), speaking rate, hesitations, and pruned syllables (i.e., a composite measure in which all dysfluencies, such as filled pauses, hesitation phenomena, and self-repetitions, are removed). Fluency varied across task types: A picture story task (the "suitcase narrative") was found to be more difficult than tasks that offered the speakers more freedom of lexical and grammatical choice, such as a monologue or a conversation. Not only did the raters perform reliably on a judgment task, but several temporal measures correlated significantly with their ratings, which suggests that the listeners were reacting similarly to fluency features present in the recorded utterances.

The second investigation in the series (Derwing, Thomson, & Munro, 2006) compared fluency ratings of both Mandarin and Slavic speakers at the start of the study, at 2 months, and at the end of the first year of residence in Canada. Although the Slavic speakers showed significantly improved fluency in terms of listener judgments over this 1-year time frame, the Mandarin speakers did not. In Derwing, Munro, and Thomson (2008), a subset of the same speakers' fluency at 2 months, 1 year, and 2 years of residence in the L2 environment was judged by another set of untrained raters to assess ongoing improvement. Again, only the ratings

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

of the Slavic speakers demonstrated a significant increase in fluency. We suggested that the differential performance of the two groups was related to overall exposure to English outside of ESL class time and that it could be understood in terms of the willingness to communicate framework proposed by MacIntyre, Clément, Dörnyei, and Noels (1998). In general, the Slavic speakers had considerably more interactions in English outside of their ESL classroom hours than did the Mandarin speakers. (See Derwing et al., 2008, for a full description of the participants' differential exposure to English.)

RESEARCH QUESTIONS

Here, we extend our previous research (Derwing et al., 2004, 2008; Derwing, Thomson, & Munro, 2006) to explore the relationship between L1 and L2 fluency. Thus, the primary focus of this article is this first research question:

1. What is the relationship between fluency ratings in the L1 and the L2 over time?

We addressed this question by collecting fluency ratings from linguistically trained speakers of the participants' L1s and from a cohort of phonetically trained English language raters who evaluated speech samples taken from the picture story task (the "suitcase narrative") at 2 months, 1 year, and 2 years. The L1 and L2 ratings were then compared statistically.

2. What is the relationship between temporal measures of fluency in the L1 and the L2 over time?

To answer this question, we measured the number of pauses per second (pause/s), the speech rate (SR; i.e., the number of syllables per second), and the pruned syllables per second (pruned syll/s). These measures were then correlated with the rating data. Additionally, because Riazantseva (2001) found that "Russian speakers with high English proficiency made significantly shorter pauses when they spoke their L2, English, than when they spoke their L1, Russian" (p. 510), it could be expected that the Russian speakers in the present study would also show changes in pause duration over time.

The data collected to address research questions 1 and 2 afforded us the opportunity to ask a third question, which has not received attention in previous studies:

3. To what extent are temporal measures of fluency in both the L1 and the L2 correlated with fluency judgments of L1 and L2 productions?

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

Although earlier work has revealed that temporal variables correlate strongly with fluency ratings in the L2 (Derwing et al., 2004; Dewaele, 1998), similar patterns of correlations for L1 data would indicate that raters are able to make reliable fluency judgments independent of accent and grammatical accuracy.

METHOD

Speakers

The speakers analyzed in this study are a subset of a larger pool of participants described in Derwing et al. (2006). Because this is a longitudinal study, there was some attrition at the 2-year point. We therefore chose to limit the number of speakers to 32 in the present study to ensure an equal number of participants in each language group. Sixteen adult native speakers of Mandarin (*M* = 34;4 years, range = 27–42; 6 men and 10 women) and 16 adult native speakers of Slavic languages (15 Russians and 1 Ukrainian; *M* = 38;6 years, range = 19–49; 8 men and 8 women) participated. All were well-educated immigrants to Canada who had arrived less than 4 months (*M* = 2;8 months, range = 1–4) before they were selected to participate. Using the Canadian Language Benchmarks Assessment tool, all participants were evaluated as beginners in listening and speaking, although most had studied English in China (*M* = 6;9 years) and the former Union of Soviet Socialist Republics (*M* = 5;7 years). The speakers were attending a full-time ESL program at the time of the first recording; by the final data collection, after 2 years, many had completed their formal ESL instruction and were either employed or enrolled in employment training programs. All speakers passed a pure-tone hearing screen (i.e., a standard test in which hearing thresholds are verified for normalcy).

Speaking Task

We sampled the Mandarin and Slavic speakers' speech on seven occasions over 2 years, using a high-quality minidisk recorder in a quiet room. Each time, the participants performed a number of tasks, including picture narrations (the "suitcase narrative" was used a total of four times). During the first data collection period (time 1), the participants examined the picture story for 1 min and then described the eight-frame set of cartoons in their L1 and in English, their L2. Only the L1 version from time 1 was used in this study. We elicited English language versions of the same set of pictures 2 months later (time 2), 10 months after the first data collection (time 6), and once more 1 year later (time 7).

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

To ensure that the L1 productions did not influence the L2 samples, we deliberately used L2 samples from times 2, 6, and 7 only rather than using the L2 production from time 1. The drawings in the picture narration task, the "suitcase narrative," depicted a man and a woman carrying identical green suitcases. They approached the same corner from different directions, accidentally bumped into each other, and fell down. They then stood up, apologized, and continued on their respective ways, only to discover later that they had mistakenly taken the wrong suitcases.

Rating Stimulus Preparation

We selected the first 30 s from the Mandarin and Slavic L1 narratives and eliminated any initial pauses. This approach helped to ensure comparable content across all speakers (Derwing et al., 2004; Munro, Derwing, & Morton, 2006). The two sets of L1 samples were randomized separately and recorded onto CDs to be played to native speakers of Mandarin and Slavic languages. Similar 20-s samples of the same speakers' L2 English productions from times 2, 6, and 7 were pooled, randomized, and recorded onto a CD for the English-speaker raters. Because of the large number of speech samples (3 samples × 32 speakers), we used shorter samples for the L2 ratings to guard against listener fatigue.

There is no consensus in the field concerning the appropriate length of speech samples for fluency evaluation by judges. In earlier studies (e.g., Derwing et al., 2006), we determined that 20 s was sufficient for raters to make a reliable assessment. It is not true that longer is necessarily better. Our decision represents a recognition of the inevitable trade-offs that occur in research of this type. Twenty seconds is subjectively perceived by raters as a long time. Using substantially longer speech samples not only lengthens the duration of the experiment, which in turn increases the demands on listeners, but also has the potential to make the rating of individual samples more difficult, because the raters must hold in memory a longer narrative. A further limitation of longer speech samples is that they may increase the likelihood of primacy and recency effects.

Native-Speaker Raters

Eight native speakers of Mandarin (two men and six women; M = 32 years) and eight native speakers of Russian (one man and seven women; M = 34 years) rated the Mandarin and Slavic productions, respectively. All of the raters were highly proficient speakers of English and all had some linguistics training. (The Russian raters were advised that one of the voices was Ukrainian. None of the raters reported difficulty judging the fluency of the Ukrainian speech samples.) Eight native speakers of

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

Canadian English (two men and six women; *M* = 42;1 years), all of whom had backgrounds in linguistics, rated the L2 productions. All raters reported normal hearing, but because the native English raters were older, on average, than the Mandarin and Russian listeners, their hearing was tested with a pure-tone screen and was found to be normal.

Rating Tasks

The ratings of the Mandarin and Slavic productions were collected individually or in pairs from the L1 listeners. We showed the participants the picture stimuli used in the narrative prior to the listening task to avoid a familiarity bias in later samples. The judges listened to the entire 30-s sample before rating each speaker on a 9-point fluency scale in the L1 (from 1 = extremely fluent to 9 = extremely dysfluent). We told the raters that we were most interested in temporal aspects of the productions (e.g., the number and length of silent pauses, filled pauses, self-repetitions, false starts). We also told them that all the speakers were educated individuals with no speech disorders and that, for these reasons, we did not expect them to use the whole scale. The judges of the English L2 speech samples also saw the picture story before completing the rating task using the same 9-point scale, and they, too, were told to focus on temporal aspects of fluency rather than general proficiency in English.

Temporal Measures

The 30-s Mandarin and Slavic speech samples were each transcribed by a native speaker of Mandarin and Russian, respectively, in standard Chinese characters and Cyrillic. The transcribers also provided character-by-character or word-for-word English glosses and overall translations as well as marks for pauses (both filled and unfilled), lexical hesitation devices used in either Mandarin or Russian, self-corrections, self-repetitions, false starts, nonlexical filled pauses, and asides. Using a waveform display program, one of the researchers then measured these phenomena to the nearest millisecond.

Similar measurements were made of the speakers' L2 productions, which were then standardized by dividing the total value of each measure by the total number of seconds of each production. We chose to use pause/s, SR, and pruned syll/s as the measures of fluency because they have been employed extensively in other studies (e.g., Lennon, 1990; Riggenbach, 1991; Towell et al., 1996). Following common practice (Derwing et al., 2004), only pauses of 400 ms or longer were considered. Once the numbers of self-corrections, self-repetitions, false starts,

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

nonlexical filled pauses, and asides were established, we subtracted all of these measures from the total number of syllables to arrive at a pruned syllables measurement, which was then standardized by calculating pruned syll/s. None of the components of the pruned syll/s measure are frequent enough to stand on their own in statistical analyses, but taken together, they result in a measurement that is a strong predictor of fluency judgments (Derwing et al.; Mehnert, 1998; Ortega, 1999).

Because informal, impressionistic assessments by the researchers suggested that the Mandarin speakers' vowel durations in English were noticeably prolonged compared with those of the other speakers, we added a further temporal measure—vowel durations of the first 12 vowels produced by each speaker in both the L1 and the L2 productions. These vowels were measured from the onset of voicing. In open syllables, the end of each vowel was defined as the point at which audible voicing concluded, whereas in closed syllables, the end point was defined as the beginning of the following consonant. For comparative purposes, we drew 16 speech samples from a database of adult native speakers of English (3 men and 13 women) who produced the same suitcase narrative. These narrators ranged in age from 24 to 46 years (M = 34 years). All reported normal hearing. Recordings were made under the same conditions and with the same equipment as for the nonnative speakers.

RESULTS

Interrater Reliability

The intraclass correlation coefficient was used to measure the degree of interrater reliability for each group of raters' evaluations of the speakers' productions. The Mandarin L1 raters' coefficient was .90, the Russian raters' coefficient was .89, and the English native-speaker raters' coefficient was .92. These results indicate high levels of agreement among all three groups of raters.

Preliminary Analyses

Prior to addressing our research questions, we conducted some preliminary analyses to verify that the data from this cohort of raters were consistent with the data obtained in our previous studies (Derwing et al., 2006, 2008). After determining consistency with the previous results, the fluency ratings were submitted to a mixed-design ANOVA with L1 group (Mandarin, Slavic) as a between-groups factor and time (2, 6, and 7) as a within-groups factor. The effect of L1 group just missed significance, $F(1, 30) = 4.17$, partial $\eta^2 = .122$, $p = .05$. However, the effect of time,

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

$F(2, 60) = 23.71$, partial $\eta^2 = .441$, $p < .001$, was significant, as was the two-way interaction, $F(2, 60) = 7.30$, partial $\eta^2 = .196$, $p < .005$. The interaction was explored by means of post hoc t tests (Bonferroni adjusted to require $p < .008$ for statistical significance), which revealed that the Mandarin group's fluency ratings did not show significant improvement between times 2 and 7, $t(15) = 2.32$, $p = .035$, whereas those of the Slavic group did, $t(15) = 6.67$, $p < .008$. However, the Mandarin group's scores between times 6 and 7 improved significantly, $t(15) = 4.53$, $p < .008$. These results generally confirm the findings of Derwing et al. (2008), which used different raters and a different scale (seven points). In particular, the Slavic speakers, but not the Mandarin speakers, were rated as more fluent at the end of the study than at the beginning. The difference in the Mandarin speakers' scores at times 6 and 7 appears to have been due to a nonsignificant tendency for that group to regress slightly from time 2 to time 6, a pattern observed in both Derwing et al. (2008) and the present study (see Figure 1). Nevertheless, no significant net change in fluency scores for the Mandarin speakers was found over the course of the study.

Individual patterns of performance within groups were also examined, as shown in Figures 2a and 2b. A comparison of the two figures shows a much larger number of learners who improved in the Slavic group ($n = 13$) than in the Mandarin group ($n = 6$). Furthermore, two of the Slavic speakers who did not improve were rated as having the best fluency at time 2; they therefore had less room to improve than the other participants. Of the eight top performers at time 2 in the Mandarin group, only one showed improvement of one point or more, whereas for the Slavic speakers, six of the top eight showed improvement.

Finally, in preparation for the parametric correlational analyses, we submitted all rating data and temporal measures to tests of normality (i.e., Shapiro-Wilks). No significant deviations from normal were observed.

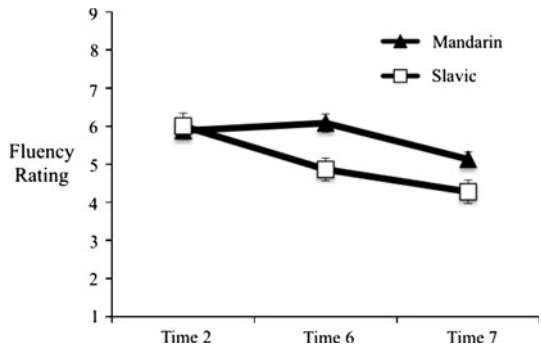


Figure 1. Mean fluency ratings (with standard error) for the Mandarin and Slavic groups at times 2, 6, and 7 (1 = extremely fluent, 9 = extremely dysfluent).

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

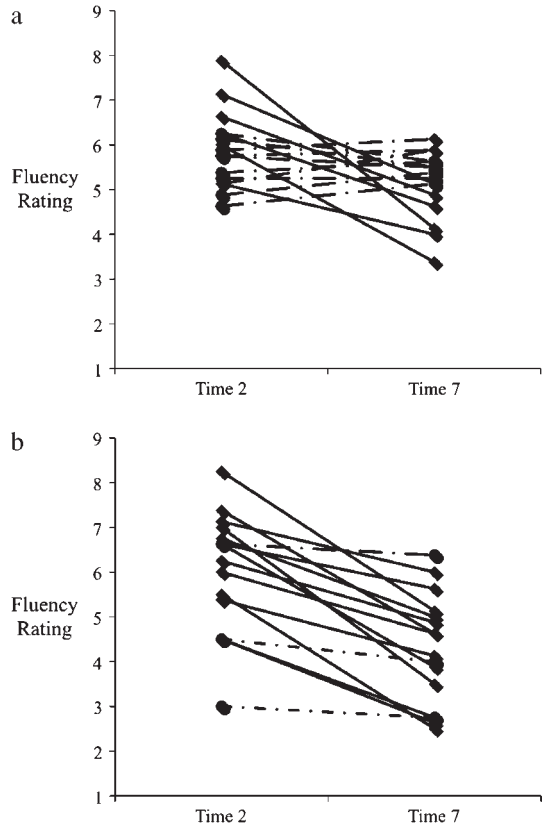


Figure 2. Changes in individual speakers' fluency ratings from time 2 to time 7. The speakers represented by solid lines improved at least one scalar unit, whereas those represented by broken lines showed either no improvement or worse ratings at time 7: (a) Mandarin group; (b) Slavic group.

Relationship Between Fluency Ratings in the L1 and the L2 Over Time

Although the data allow a correlational examination of L1 and L2 fluency, they cannot be used to compare L1 and L2 fluency directly because different raters were used and the variability within the L1 and the L2 speaker groups was different. In the L1 condition, for example, the listeners' reference points for highly fluent versus dysfluent speech were likely quite different from those of the listeners who heard the L2 samples. Thus, the range for the listeners of fluent versus dysfluent speech was somewhat smaller.

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

Pearson correlations between the fluency ratings of the same speakers' L1 and L2 indicated that there was a significant positive correlation for both Mandarin- and Slavic-speaking participants at time 2 (i.e., when they were still relatively beginning learners) but that, at times 6 and 7, there were no significant relationships, as shown in Table 1. The lack of significance at times 6 and 7 cannot simply be due to reduced variability in fluency scores at those times or to insufficient variability in the L1 scores. In fact, the fluency scores at time 6 showed a wider span for the Mandarin speakers than at time 2. Moreover, the fact that L1 fluency scores correlated significantly with time 2 fluency ratings demonstrates that there was sufficient variability in the L1 ratings. Scatterplots that show the scores for individual speakers comparing L1 fluency ratings with L2 ratings at times 2 and 7 are provided in Figures 3a, 3b, 4a, and 4b.

To determine whether L1 fluency ratings were predictive of the speakers' degree of improvement over time, we carried out an additional analysis in which we compared the fluency scores in the L1 with the change in fluency scores between time 2 and time 6 and between time 2 and time 7. Pearson correlations (df =14) were not significant, either for the Mandarin speakers, r = –.454, p = .08 and r = –.496, p = .05, respectively, or for the Slavic speakers, r = –.264, p = .32 and r = –.336, p = .20, respectively. Thus, there was no statistically significant indication that L1 fluency scores were related to the speakers' improvement in L2 fluency over time, despite an initial relationship at time 2.

Relationship Between Temporal Measures of Fluency in the L1 and the L2 Over Time

When the L1 measures of pause/s, SR, and pruned syll/s were compared with the same measures for the L2, there was a significant correlation at time 2 for both groups, similar to the relationship observed for the

Table 1. Correlations between L1 and L2 fluency ratings

	Mandarin	Slavic
Time	Pearson's r	Pearson's r
2	.527*	.620*
6	–.043	.477
7	–.253	.401

*p < .05.

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

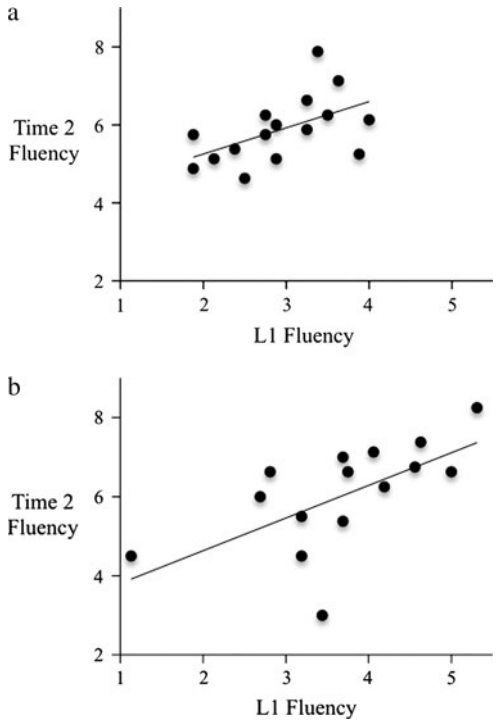


Figure 3. Relationship between L1 fluency and L2 fluency at time 2: (a) Mandarin group; (b) Slavic group.

fluency ratings. Figures 5a–5d present the mean values for each of the temporal measures (i.e., pruned syll/s, SR, pause/s, and pause length, respectively), and Table 2 presents the results of the correlational analyses. At times 6 and 7, significant L1 and L2 correlations were found for the Slavic group only (with the exception of pruned syll/s at time 7; see the Appendix for fluency ratings and temporal measurements).

Although the temporal properties of the speakers' utterances have been evaluated in Derwing et al. (2004), we carried out one additional analysis to address a point arising from Riazantseva's (2001) research. To determine whether her finding of differential pausing over time held for the data under consideration here, we computed one-way repeated measures ANOVAs for each speaking group and compared the mean durations of unfilled pauses in the L1 and at times 2, 6, and 7 (see Figure 5d).

For the Mandarin speakers, the main effect was nonsignificant, $F(3, 45) = 2.361$, $p = .084$, and for the Slavic speakers, the main effect was also nonsignificant, $F(3, 45) = 0.259$, $p = .855$. Thus, there was no indication of any difference between mean pause times in the L1 and

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

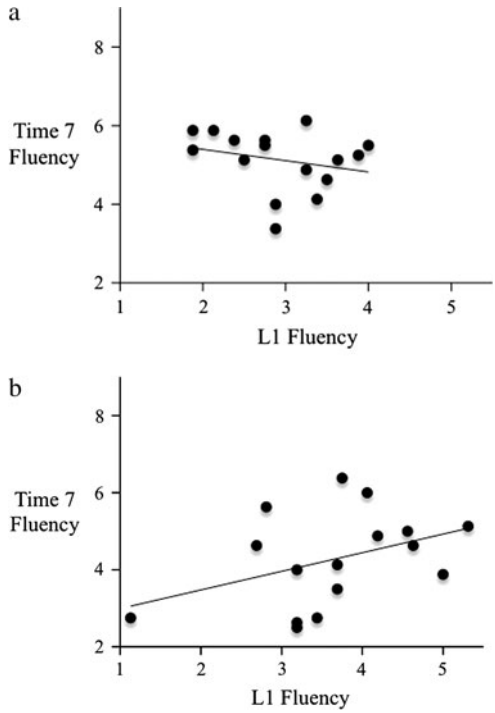


Figure 4. Relationship between L1 fluency and L2 fluency at time 7: (a) Mandarin group; (b) Slavic group.

pause times at any of the data collection points in the L2 for either language. Nor did the speakers differ across the three L2 data collection times.

As indicated in Figure 6, the mean vowel durations for the first 12 vowels in each narrative were 152, 107, and 108 ms for Mandarin, Slavic, and English, respectively. A one-way between-groups ANOVA with three levels of L1 that compared durations across the three languages yielded a significant main effect, $F(2, 45) = 9.889$, $p < .001$. Post hoc t tests (Bonferroni-adjusted to require $p < .017$) indicated that the Mandarin mean vowel duration was significantly longer than the mean durations of both the Slavic, $t(30) = 3.7715$, $p < .001$, and English, $t(30) = 3.3675$, $p < .005$, vowels. However, the difference between Slavic and English durations was not significant, $t(30) = 0.0472$, $p = .9627$.

An additional analysis was carried out on the English vowel productions of both groups of L2 speakers. A mixed-design ANOVA with L1 group (Mandarin, Slavic) as a between-groups factor and time (2, 6, and 7) as a within-groups factor yielded a significant effect of L1, $F(1, 30) = 20.791$, partial $\eta^2 = .409$, $p < .001$, and of time, $F(2, 60) = 5.091$, partial $\eta^2 = .145$,

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

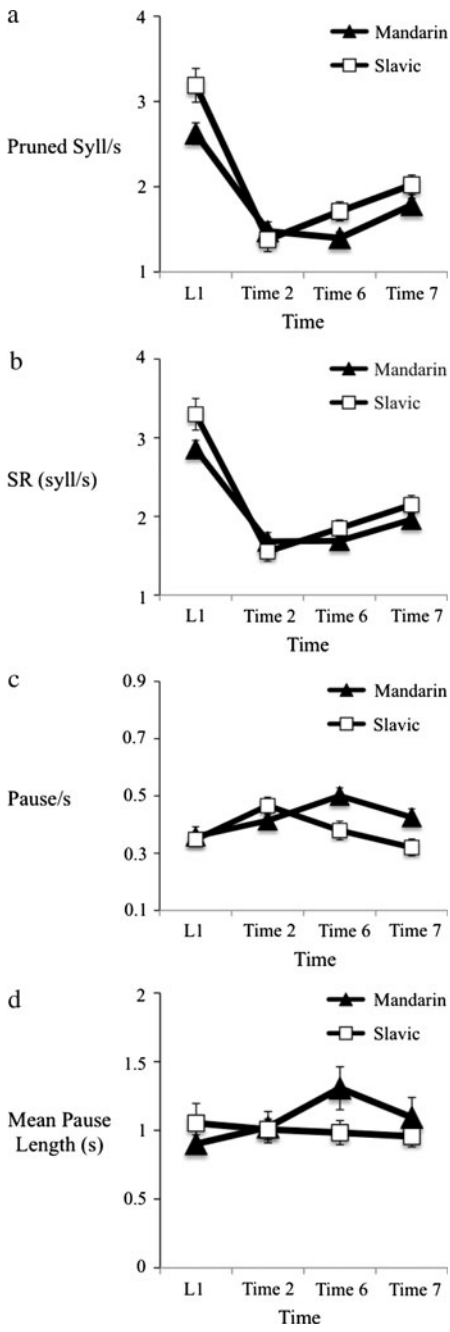


Figure 5. Mean (a) pruned syll/s, (b) SR, (c) pause/s, and (d) pause length, for the Mandarin and Slavic groups in the L1 and in the L2 at times 2, 6, and 7. Error bars represent standard error.

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

Table 2. Pearson correlations of L1 and L2 temporal measures

Temporal measure	Mandarin	Slavic
Pause/s		
2	.524*	.722**
6	–.050	.599*
7	.400	.569*
SR (syll/s)		
2	.618*	.583*
6	–.090	.613*
7	.388	.548*
Pruned syll/s		
2	.556*	.638**
6	–.286	.625**
7	.279	.488

Note. All temporal measures were standardized by dividing by the total number of seconds.
*p < .05 (two-tailed).
**p < .01 (two-tailed).

p < .01. The two-way interaction was not significant, $F(2, 60) = 0.470$, partial η^2 = .015, p = .627. These results indicate that the Mandarin speakers' English vowels were longer than those of the Slavic speakers and that the English vowels in both groups became significantly shorter over time.

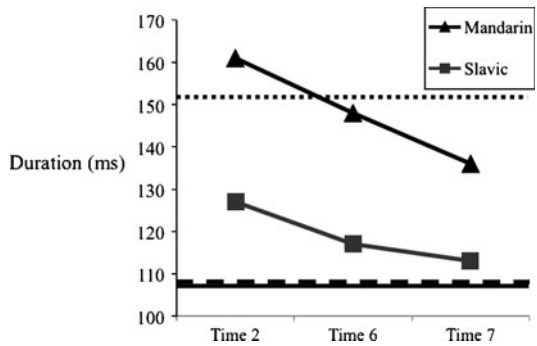


Figure 6. Mean L2 English vowel durations over time are contrasted with mean L1 vowel durations. The solid line indicates mean L1 Slavic vowel durations, the dashed line indicates mean L1 English vowel durations, and the dotted line indicates mean L1 Mandarin vowel durations.

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

Correlations of Temporal Measures of Fluency With L1 and L2 Productions

Pearson correlations between the L1 fluency ratings and temporal measures (i.e., L1 pause/s, SR, and pruned syll/s) as well as correlations between L2 fluency ratings and these same temporal measures at each of the three collection times are provided in Table 3. Fluency scores were significantly correlated with at least one of the three temporal measures for both groups in the L1 and at all times in the L2; however, the correlations were almost always stronger for the Slavic speakers.

To determine whether the L2 vowel durations and the fluency ratings were related, Pearson correlations between the two sets of measures for all three times were computed (see Table 4). No significant correlations were found.

Table 3. Pearson correlations between fluency ratings and temporal measures

		Time		
Fluency ratings	L1	2	6	7
Mandarin				
Pause/s	.395	.703**	.125	.798**
SR (syll/s)	–.519*	–.503*	–.261	–.578*
Pruned syll/s	–.662**	–.654**	–.533*	–.420
Slavic				
Pause/s	.573*	.697**	.444	.384
SR (syll/s)	–.764**	–.850**	–.827**	–.741**
Pruned syll/s	–.847**	–.881**	–.898**	–.810**

Note. All temporal measures were standardized by dividing by the total number of seconds.
*$p < .05$ (two-tailed).
**$p < .01$ (two-tailed).

Table 4. Correlations between mean L2 vowel durations and mean fluency ratings

	Time		
Group	2	6	7
Mandarin	.099	–.128	–.074
Slavic	.034	.406	.488

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

DISCUSSION

In this study, we examined fluency ratings and temporal measures of the L2 with comparable measures in the L1 to determine whether speakers who were extremely fluent in their L1 would be more fluent than their peers in the L2. Is fluency a trait that transfers across languages, or is it a state that must be developed independently? The clearest relationship between L1 and L2 fluency ratings was observed in the initial stages of L2 exposure, and, overall, this relationship was stronger in the Slavic than in the Mandarin learners of English. However, at later stages, the correlation between the judgments was lower and not statistically significant in either group. Nor was there a statistical relationship between L1 fluency and improvement in fluency over time. These findings indicate that the state versus trait question cannot be fully resolved with this study. On the one hand, our results are not very consistent with a trait view, at least in the first 2 years of language learning. If fluency were strictly a trait, applicable in both the L1 and the L2, we would expect a continued relationship between L1 and L2 fluency over time, provided that the learners gained increasing proficiency in their L2 at the same rate. For the participants observed here, there was a weakening of the relationship between L1 and L2 fluency ratings over time, but there was also a difference in degree of voluntary exposure to English between the two groups (Derwing et al., 2008). It is not possible to determine whether the weakening relationship was due to increased variability in English proficiency across the learners or to the nontraitlike nature of fluency. On the other hand, if fluency were strictly a state, rather than a trait, the fluency of highly proficient L2 speakers should vary in ways that are not predictable from L1 fluency performance. In other words, advanced speakers should be more fluent in their L2 than in their L1 just as often as the reverse, holding contextual factors such as topic, interlocutor, and setting constant. Because we have not examined L2 speakers at the stage of ultimate attainment, it is not possible for us to test this prediction.

The discrepancy in the learners' fluency ratings at different collection times raises an important methodological issue. Had we relied solely on a comparison of L1 and L2 fluency ratings at time 2, we might have drawn the conclusion that there is a straightforward relationship between L1 and L2 fluency. However, the lack of correlation at time 7, nearly 2 years after time 2, indicates that such a conclusion would have been unwarranted. This finding points to the necessity of ongoing longitudinal research to achieve a full understanding of L2 fluency development. For instance, many of the Mandarin learners in this study may have not yet crossed a necessary threshold that would allow them to make significant progress in fluency development because of their more limited exposure to and experience with English compared with the Slavic speakers over the course of this study (Derwing et al., 2008).

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

When we examined temporal measures (i.e., pause/s, SR, and pruned syll/s), again a relationship was found between the L1 and the L2 in both groups at time 2 for all three measures, but only the Slavic group showed significant correlations at times 6 and 7, and these were weaker than at time 2. With respect to SR, this outcome is similar to that of Towell and Dewaele (2005), who found significant correlations between L1 and L2 SRs for learners of French, although the strength of the correlation was slightly lower at time 2 of their study than at time 1. The reason for this patterning is unclear, but it may be that, especially in the early stages of learning, the Slavic speakers benefited from the structural parallels between their L1 and English. In the case of the Mandarin speakers, perhaps the most effective language learners were those whose L1 fluency scores were best. These individuals may have already capitalized as much as possible on their superior oral skills by time 2 in the study. The lack of a relationship between L1 and L2 fluency at the later times appears to be due to a lack of improvement in these individuals, whereas the other participants started to catch up.

We evaluated mean unfilled pause durations to test Riazantseva's (2001) contention that pause durations for Russian speakers would decline in English as their proficiency increased. Although Riazantseva found that high-proficiency Russian speakers of English paused less in their L2 than in their L1, we were unable to replicate this finding. The time 7 unfilled pause durations for Slavic speakers (many of whom were advanced speakers by that point) were not significantly different from their L1 unfilled pauses and pause durations and did not decline between the end of the first year (time 6) and the end of the second year (time 7). It is not clear why this discrepancy in findings occurred.

Our examination of one additional temporal measure—vowel duration—brought to light another difference between the two participant groups: The Slavic speakers eventually approached nativelike English vowel durations more closely than the Mandarin speakers. Although both groups showed improvement, the closer approximation by the Slavic speakers may be attributed to transfer from the L1, given that Russian has vowel reduction much like English but Mandarin does not. Transfer from the L1 may also be implicated in the case of the Mandarin speakers, whose L1 vowels may be longer on average than some English vowels because they carry tone (Zhang, 2002). Longer vowel durations in the Mandarin productions may have led the English listeners to assign harsher ratings because of their expectations about the role of vowel duration in stress assignment (Hahn, 2004; Zielinski, 2008).

The comparison of temporal measures and L2 fluency ratings at times 2, 6, and 7 confirmed that temporal measures of the type typically used in fluency research correlate with listener judgments. However, the listeners appeared to assign ratings that better reflected temporal measures for the Slavic than for the Mandarin speakers. We can only

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

speculate that there were other features in the Mandarin speech that influenced the listeners' ratings. Most importantly, however, the temporal variable that best predicted L1 fluency (i.e., pruned syll/s) also emerged as the strongest predictor of most of the L2 fluency ratings. This occurred even though the L1 and L2 raters were different groups of participants. Methodologically, this finding attests to the meaningfulness of listener judgments of both L1 and L2 speech. It also suggests that raters of L2 speech can indeed focus on fluency even when other aspects of the speech stream (e.g., accentedness and grammatical inaccuracies) are present.

CONCLUSIONS

These findings offer valuable insights into the complex relationship between L1 and L2 fluency—namely, that a straightforward relationship between fluency in the L1 and the L2 cannot be expected. As Segalowitz and Freed (2004) observed, easily interpretable results are virtually nonexistent. A variety of factors, including the structural properties of the L1 itself, the proficiency levels of the speakers, the degree of exposure to the L2 (including instruction), and cognitive factors such as working memory (Segalowitz et al., 1998), all contribute to fluency development. Resolving the state versus trait issue will require further investigations of these areas and may entail individual cognitive assessments of adult L2 learners.

These results also point to pedagogical implications for the English language classroom. The Mandarin learners, and, to a lesser extent, the Slavic learners, might well have benefited from direct fluency instruction such as the activities advocated by Nation (1989) and Gatbonton and Segalowitz (2005) because, despite living in an almost monolingual English community, the participants did not have access to many opportunities to speak English outside of class. Their ESL program focused heavily on the development of written language and academic English, with little time left for oral skills practice, which perhaps resulted in an increase in declarative knowledge with little or no concomitant enhancement of procedural knowledge. In interviews, several participants reported that they would have liked more listening and speaking practice in their ESL classes. Perhaps there was an assumption on the part of the program organizers that the students would be able to supplement their classroom experiences in the larger community, although for many learners that was not the case. To help learners take advantage of fluency development opportunities, ESL programs could plan contact activities that would benefit both the learners and the members of the larger community with whom they would interact. Dudley (2007) provided suggestions for volunteering experiences that could lead to greater oral fluency and proficiency in adult L2 learners.

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

A limitation of the present study is that it is restricted to the first 2 years of linguistic development in adult immigrant learners, who had not yet reached a state of ultimate attainment. More extensive longitudinal investigations are necessary to establish the long-term relationship between L1 and L2 fluency. Ortega and Iberri-Shea (2005) have observed the need for much longer studies to cover the time frame during which significant aspects of L2 development occur. Furthermore, given the quite different patterns of performance between the two groups who participated here, investigations that involve a greater range of L1s should be carried out. It may be that there are closer L1 and L2 fluency relationships in some language pairings than in others. Further insights may also be gained by a much closer examination of individual learner performance. Finally, it will only be with more research that we can discover why Ellis and Larsen-Freeman's (2006) observation that "the very things that make a known language easy make a new language hard" (p. 568) seems to apply to some, although not all, participants in this study.

(*Received 5 December 2008*)

REFERENCES

Chambers, F. (1997). What do we mean by fluency? *System*, *25*, 535–544.
Collentine, J., & Freed, B. F. (2004). Learning context and its effects on second language acquisition. *Studies in Second Language Acquisition*, *26*, 153–171.
de Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics*, *13*, 1–24.
Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, *29*, 359–380.
Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). L2 fluency: Judgments on different tasks. *Language Learning*, *54*, 655–679.
Derwing, T. M., Thomson, R. I., & Munro, M. J. (2006). English pronunciation and fluency development in Mandarin and Slavic speakers. *System*, *34*, 183–193.
Dewaele, J.-M. (1998). Speech rate variation in two oral styles of advanced French interlanguage. In V. Regan (Ed.), *Contemporary approaches to second language acquisition in social context* (pp. 113–123). Dublin, Ireland: University College Academic Press.
Dewaele, J.-M. (2002). Individual differences in L2 fluency: The effect of neurobiological correlates. In V. Cook (Ed.), *Portraits of the L2 user* (pp. 221–249). Clevedon, UK: Multilingual Matters.
Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, *26*, 249–273.
Dudley, L. (2007). Connecting L2 learners to the larger community. *Canadian Modern Language Review*, *63*, 539–561.
Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, *27*, 164–194.
Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics. *Applied Linguistics*, *27*, 558–589.
Fillmore, C. J. (1979). On fluency. In C. J. Fillmore, D. Kemple, & W. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85–101). San Diego, CA: Academic Press.
Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French. *Studies in Second Language Acquisition*, *26*, 275–301.

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review*, *61*, 325–353.
Guillot, M.-N. (1999). *Fluency and its teaching*. Clevedon, UK: Multilingual Matters.
Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*, 201–223.
Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.
Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Erlbaum.
Lafford, B. A. (2004). The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition*, *26*, 201–225.
Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40*, 387–413.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.
MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journal*, *82*, 545–562.
Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, *20*, 83–108.
Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition*, *23*, 451–468.
Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, *28*, 111–131.
Nation, I. S. P. (1989). Improving speaking fluency. *System*, *17*, 377–384.
O'Brien, I., Segalowitz, N., Freed, B. F., & Collentine, J. (2007). Phonological memory predicts second language oral fluency in adults. *Studies in Second Language Acquisition*, *29*, 557–581.
Ortega, L. (1999). Planning and focus on form in L2 performance. *Studies in Second Language Acquisition*, *21*, 109–148.
Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, *25*, 26–45.
Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: Benjamins.
Raupach, M. (1980). Temporal variables in first and second language speech production. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 263–270). Berlin: Mouton de Gruyter.
Riazantseva, A. (2001). Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition*, *23*, 497–526.
Riggenbach, H. (Ed.). (2000). *Perspectives on fluency*. Ann Arbor: University of Michigan Press.
Riggenbach, H. (1991). Towards an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, *14*, 423–441.
Segalowitz, N. (2007). Access fluidity, attention control, and the acquisition of fluency in a second language. *TESOL Quarterly*, *41*, 181–186.
Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, *26*, 173–199.
Segalowitz, N., Freed, B. F., Collentine, J., Lafford, B. A., Lazar, N., & Díaz-Campos, M. (2004). A comparison of Spanish second language acquisition in two different learning contexts: Study abroad and the domestic classroom. *Frontiers, 10*, 1–18.
Segalowitz, N., & Hulstijn, J. H. (2005). Automaticity and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–378). Oxford: Oxford University Press.
Segalowitz, S. J., Segalowitz, N., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics*, *19*, 53–67.

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49*, 93–120.
Towell, R. (2002). Relative degrees of fluency: A comparative case study of advanced learners of French. *International Review of Applied Linguistics*, *40*, 117–150.
Towell, R., & Dewaele, J.-M. (2005). The role of psycholinguistic factors in the development of fluency. In J.-M. Dewaele (Ed.), *Focus on French as a foreign language* (pp. 210–239). Clevedon, UK: Multilingual Matters.
Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, *17*, 84–119.
Trofimovich, P., & Baker, W. (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.
Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*, 105–122.
Wood, D. (2001). In search of fluency: What is it and how can we teach it? *Canadian Modern Language Review*, *57*, 573–589.
Zhang, J. (2002). *The effects of duration and sonority on contour tone distribution: A typological survey*. London: Routledge.
Zielinski, B. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, *36*, 69–84.

APPENDIX

MEAN FLUENCY RATINGS AND TEMPORAL MEASUREMENTS

	Mandarin speakers		Slavic speakers	
Measure	*M*	*SD*	*M*	*SD*
Fluency rating				
L1	2.9	0.67	3.7	1.02
Time 2	5.9	0.85	6.0	1.35
Time 6	6.1	0.95	4.9	1.19
Time 7	5.1	0.76	4.3	1.22
Pause/s				
L1	0.4	0.09	0.3	0.14
Time 2	0.4	0.12	0.5	0.13
Time 6	0.5	0.13	0.4	0.11
Time 7	0.4	0.12	0.3	0.12
SR				
L1	2.9	0.44	3.3	0.80
Time 2	1.7	0.45	1.6	0.49
Time 6	1.7	0.36	1.8	0.42
Time 7	2.0	0.31	2.1	0.47
Pruned syll/s				
L1	2.6	0.50	3.2	0.79
Time 2	1.5	0.43	1.4	0.56
Time 6	1.4	0.29	1.7	0.45
Time 7	1.8	0.32	2.0	0.46

https://doi.org/10.1017/S0272263109990015 Published online by Cambridge University Press