Article

LANGUAGE
TEACHING
RESEARCH

Language Teaching Research
1–24

DOI: 10.1177/13621688221076418
journals.sagepub.com/home/ltr
SAGE

Investigating the contribution of L1 fluency, L2 initial fluency, working memory and phonological memory to L2 fluency development

Nancy Gagné
Université TÉLUQ, Canada

Leif M. French
Sam Houston State University, USA

Kirsten M. Hummel
Université Laval, Canada

Abstract

Within the same learning context, learners' outcomes in terms of oral fluency vary greatly. This study tracked the relative contributions that first language (L1) and initial second language (L2) fluency skill and working memory (WM) made to L2 fluency development. We assessed the performance of French-speaking Grade 6 learners' (*n* = 47, mean age: 11) in a 10-month intensive English program in Quebec, Canada using a picture-cue monologic task based on The Suitcase Story and a semi-structured interview based on the Oral Proficiency Interview (OPI). Working memory was assessed using a backward digit span task and phonological memory (PM) via non-word repetition and serial non-word recognition tasks. Overall, results suggest that L1 fluency, WM and PM played only a minor role in learners' L2 fluency outcomes, whereas learners' pre-program levels of L2 fluency constituted an important predictor of L2 fluency development regardless of the speech task used to index fluency.

Keywords

fluency, intensive instruction, memory, second language learning, young learners

---

Corresponding author:
Nancy Gagné, Département Sciences Humaines, Lettres et Communication, Université TÉLUQ, 455, rue du Parvis, Québec, QC G1K 9H6, Canada
Email: nancy.gagne@teluq.ca

I Introduction

Researchers of second language acquisition (SLA) have frequently observed that most learners rarely reach the same fluency level in their second language (L2) as in their first language (L1) (Birdsong, 2005, 2006, 2009, 2014; De Jong et al., 2015; Derwing et al., 2009; Duran-Karaoz & Tavakoli, 2020; Hilton, 2008a, 2008b; Hulstijn & Bossers, 1992). In these studies, researchers have identified a number of factors that may explain this variation, such as learning context (type, duration, intensity of L2 exposure), willingness to communicate, communication anxiety or behavioral inhibition, age of onset, or aptitude (Ellis & Ferreira, 2009; Foster, 2020; Muñoz & Singleton, 2011; Segalowitz & Hulstijn, 2005; Zabihi, Ghominejad & Javad Ahmadian, 2021). However, other research has increasingly suggested that individual differences in L1 fluency may influence L2 fluency (De Jong et al., 2015; Derwing et al., 2009; Peltonen, 2018; Segalowitz, 2010; Towell & Dewaele, 2005) and that traits common to both L1 and L2 may explain individual variability in L2 fluency (De Jong & Mora, 2017; Segalowitz, 2010, 2016). Further, individual differences in cognitive abilities (e.g. lexical access, attention or working memory) may also partially explain L2 fluency attainment (Linck et al., 2014; O'Brien et al., 2007). However, to date, little research has examined the independent and combined influences of L1 fluency, initial L2 fluency level, and cognitive abilities on the development of L2 fluency. To address this issue, we examined working memory (backward digit span), phonological memory (non-word repetition, serial non-word recognition) and both L1 and L2 fluency (temporal measures extracted from monologic and dialogic tasks) in a group of EFL learners enrolled in a 9-month intensive English program.

II L1 and L2 fluency

Segalowitz (2010) provides a well-documented theoretical model to study L2 fluency. Within this largely cognitive model, fluency is defined as, 'how efficiently the speaker is able to mobilize and temporarily integrate, in a nearly simultaneous way, the underlying processes of planning and assembling an utterance in order to perform a communicatively acceptable speech act' (p. 47). The model defines three specific fluency notions: utterance fluency, cognitive fluency, and perceived fluency. In the present study, we focus on utterance fluency, which includes objective temporal measures of speech production (e.g. speech rate) (Chambers, 1997; Cucchiarini, Strik & Boves, 2000, 2002; Lennon, 1990; Riggenbach, 1991; Towell & Dewaele, 2005; Towell, Hawkins & Bazergui, 1996). Like Housen et al. (2012) we consider fluency a phonological component and argue that temporal measures of utterance fluency provide insight into the cognitive processes related to language processing (Linck et al., 2014).

Previous research has highlighted initial L2 fluency level as an individual predictor of L2 outcomes (Segalowitz & Freed, 2004; Valls-Ferrer & Mora, 2014) and, with increasing frequency, L1 fluency as a predictor of L2 fluency development (De Jong et al., 2015; Derwing et al., 2009; Duran-Karaoz & Tavakoli, 2020; Peltonen, 2018; Segalowitz, 2010; Towell & Dewaele, 2005). The majority of L2 fluency studies have assed various objective temporal measures – e.g. speech rate (SR), phonation time (PT), or mean length of run (MLR) – in either picture description tasks (e.g. Derwing, Munro & Thomson, 2008; Derwing et al., 2009; Munro & Derwing, 1995) or dialogic tasks (O'Brien et al., 2007).

Several studies have identified a relationship between L1 and L2 fluency using monologic tasks. In a 4-year longitudinal study of adult English-speaking students who studied abroad in France, Towell and Dewaele (2005) found a positive correlation between L1 and L2 fluency on SR at Time 1 (T1) and at T2. Derwing et al. (2009) also studied the relationship between fluency in L1 and L2 in a 2-year longitudinal study with adult learners of English who had typologically distinct L1s (Chinese and Russian). They found correlations between L1 temporal measures and the same L2 measures at T2. Interestingly, this relationship disappeared by T6 and T7 for some learners, suggesting that the strength of the relation between the L1 and the L2 may decrease with increased proficiency. Similarly, Riazantseva (2001) found that intermediate adult learners' pause patterns resembled their L1 (Russian), while more advanced adult learners' production paralleled L2 (US English) pause duration. This suggests that L2 learners may better inhibit certain aspects of L1 transfer at more advanced levels.

Researchers have also used picture narrative tasks to examine whether L1 fluency relates to L2 fluency. Huensch and Tracy-Ventura (2017) examined the extent to which L1 (English) fluency predicted L2 (Spanish) fluency in adult learners who participated in a 5-month study abroad program. Regression analyses demonstrated that L1 fluency measures contributed significantly to variation in L2 fluency development. The authors also showed that crosslinguistic differences in temporal measures of L1 fluency may predict L2 fluency. More recently, Duran-Karaoz & Tavakoli (2020), investigated whether different proficiency levels (i.e. A2, B1 and B2) mediated the relationship between L1 (Turkish) and L2 (English) fluency in adult learners. They found that, after controlling for proficiency levels, the relationship between L1 and L2 fluency was maintained, suggesting that the relationship was not mediated by proficiency level. They also found that L1 fluency behaviors such as mid-clause filled and mid-clause silent pauses, the total number of repairs and speech rate predicted L2 fluency behaviors, with SR in the L1 being the best predictor of L2 fluency behaviors.

Together, these findings suggest a relationship between L1 and L2 temporal fluency measures. However, methodological differences in fluency assessment and task type(s) may confound cross-study comparisons. In fact, previous research has rarely considered task type and related constraint(s) on L2 speech production (Préfontaine & Kormos, 2015). In a study of Cantonese- and Mandarin-speaking adult learners of English, Skehan, Foster and Shum (2016) examined task effect on L2 speech production using four different kinds of monologic tasks (varying from non-structured production to a highly predictable structured task) and found that more structured (i.e. predictable) tasks led to better temporal measures (e.g. longer runs). The authors suggested that learners were able to more easily connect and scaffold ideas in structured monologic tasks, thus increasing temporal fluency measures. In a study of adult L1 English speakers enrolled in a French immersion program, Préfontaine and Kormos (2015) have demonstrated that speakers tend to be more fluent in dialogic tasks than in monologic tasks. To our knowledge, no research has thus far assessed the effects of tasks on fluency development in an intensive learning context.

III Individual differences in working and phonological memory

Variations in L2 fluency attainment could also be related to differences in cognitive abilities, especially those related to memory. Researchers have attempted to define working memory (WM) using various models (Cowan, 1995; Engle et al., 1999; Kieras et al., 1999). Baddeley and Hitch's (1974) multi-component model of WM is the most commonly used in second language acquisition research (Juffs & Harrington, 2011). In this model, WM is viewed as a system responsible for actively and temporarily storing information required to execute complex cognitive tasks such as comprehension of verbal material (Baddeley, 2003). This complex, interactive system plays an integral role in linking cognition and action needed to manipulate information from different sources (e.g. sounds and images) and at different stages of information processing (Baddeley, 2012; Gathercole & Baddeley, 2014). A central executive component provides attention control and coordination to focus, divide and switch attention. It interacts with three slave systems: the visual sketchpad, the episodic buffer (a multimodal information storage system) and the phonological loop, the latter consisting of a phonological store and an auditory rehearsal unit responsible for maintaining and manipulating verbal information (Baddeley, 2000, 2003, 2012).

Traditionally, researchers have assessed WM using span tasks, such as forward digit recall and non-word list recall, which focus on the storage component of WM. WM is also measured by tests in which participants simultaneously perform two tasks, such as information recall (storage component) and repeating a sequence in inverse order (processing component). Research suggests that complex tasks, such as backward digit recall (Marini et al., 2019), target the central executive, which relates to speech production (Kormos & Sáfár, 2008; Kudo & Lee Swanson, 2014).

Regardless of learning context there has been consensus for Ellis' (2001) hypothesis that WM appears to be one of the best predictors of success in any learning situation (Brooks, Kempe & Sionov, 2006; Ellis, 2001; Erlam, 2005; Robinson, 2005). Researchers have investigated the relationship between WM and L2 speech production and found clear evidence for a link (Fortkamp, 1998, 1999, 2000; Georgiadou & Roehr-Brackin, 2017; Gilabert & Muñoz, 2010; Kormos & Sáfár, 2008; Mizera, 2006; Trebits & Kormos, 2008). Because the speaker has to maintain parts of the message in memory while simultaneously planning subsequent parts of speech (Kormos, 2006), even small variations in WM might result in different levels of success in L2 learning (Kormos & Sáfár, 2008).

As a WM subsystem, phonological memory (PM) has also consistently shown a strong link to language acquisition. Since PM is responsible for processing verbal material, and speech production itself relies heavily on PM operations (Baddeley, 2000, 2003; Baddeley, Gathercole & Papagno, 1998; Baddeley & Hitch, 1974; Vallar & Baddeley, 1984a, 1984b), PM likely has an impact on the degree of success in L2 learning (Kormos & Sáfár, 2008).

PM's capacity has typically been measured by word, non-word, or digit recall tests. Second language acquisition studies have frequently used non-word repetition (NWR) tasks to study the relationship between PM and L2 competency in both children and adults (Abdallah, 2010; French & O'Brien, 2008; Hummel, 2009; Hummel & French, 2016; Lord, 2006; O'Brien et al., 2006; Vulchanova et al., 2014). Repeating non-words aloud immediately after hearing them engages the rehearsal process (Poncelet & Van der Linden, 2003), which reduces reliance on long-term lexical knowledge to support the temporary storage of phonological forms in short-term memory (Gathercole & Baddeley, 1989). However, although the NWR task has been used successfully in different study contexts, some researchers have argued that these tasks contain an articulation component, therefore potentially confounding articulation with PM skill (Gathercole et al., 2001; O'Brien et al., 2006).

To mitigate the influence of articulatory mechanisms on NWR task performance, researchers have often used a serial non-word recognition (SNWR) task (O'Brien et al., 2006). Participants who complete this task (comparing non-word series for similarity) are unable to rely on pre-existing lexical representations that could lead to lexicality effects (Gathercole et al., 1991, 1999, 2001; Hulme, Maughan & Brown, 1991). This task provides a more specific measure of PM by isolating the capacity of the phonological store and the rehearsal process and has also been shown to be related to L2 fluency development (O'Brien et al., 2007).

Research exploring the relationship between WM and PM in L2 speech production has shown links between these variables and different skills (e.g. reading, listening, and speech production (Fortkamp, 1998, 1999, 2000; French & O'Brien, 2008; Hummel & French, 2016; Kormos & Sáfár, 2008; O'Brien et al., 2006, 2007). In a large-scale meta-analysis, Linck et al. (2014) highlighted the robustness of the relation between WM and L2 development, showing that WM and PM may be involved in different aspects of L2 learning such as vocabulary, grammar, and oral production (e.g. French & O'Brien, 2008; Kormos & Sáfár, 2008; O'Brien et al., 2006, 2007). PM may also be related to fluency development, particularly with learners at early stages of acquisition. (O'Brien et al., 2007).

To date, very few studies have assessed both WM and PM in an L2 developmental context. In their study of adolescent L1 Hungarian learners of English, Kormos and Sáfár (2008) found that WM (L2 backward digit span, or BDS) and PM (NWR) were positively correlated to measures of reading and listening comprehension and oral production. Kormos and Sáfár's findings align with previous research with L1 children (Adams & Gathercole, 1995, 2000). These studies showed that high scores on digit span tasks were associated with longer sentences, greater syllable diversity, and a more extensive lexical repertoire, suggesting a specific link between PM and oral fluency. Moreover, Kormos and Sáfár (2008) found a moderate correlation between PM and fluency in pre-intermediate learners, but no such link in beginners. This contradicted previous findings that showed a diminishing correlation between subjective fluency measures and PM as proficiency increased (Cheung, 1996; Gathercole et al., 2005; Speciale, Ellis & Bywater, 2004).

Using monologic tasks and tasks targeting WM (speaking span and math span) and PM (NWR), Mizera (2006) investigated the link between WM, PM and L2 fluency. The adult study population included English language learners from a variety of linguistic backgrounds, as well as English-speaking learners of Spanish. He found a weak correlation between SR and speaking span, with slightly higher correlations in the lower-proficiency group than among higher-proficiency learners. Although somewhat contradictory to previous studies, perhaps due to methodological differences, these results suggest that, at the very least, WM may potentially interact with proficiency to influence speaking skill.

More recently, Georgiadou and Roehr-Brackin (2017) investigated the relationship between fluency and WM (L1 BDS and L2 listening span) and PM (recall task) in two different proficiency groups of L1 Arabic learners of English. They reported no significant relationship between WM measures and fluency for beginning-level participants. However, they found a significant negative correlation between BDS and the number of pauses in the lower intermediate group, revealing that participants with greater executive WM paused less. However, assessing WM in the L2 might have been a confounding factor in the study and could explain why no specific links were found in beginners.

Little research has specifically examined the relationship between PM and L2 fluency. A notable exception is O'Brien et al.'s (2007) groundbreaking study that investigated PM's relative contribution to L2 fluency development for English-speaking learners of Spanish over time two contexts (i.e. instructional and study abroad). Using a SNWR task to measure PM and an Oral Proficiency Interview (OPI) (ACTFL, 2012) to evaluate fluency, they found that participants improved on 5 of 6 fluency measures (defined in terms of speech measures and oral fluidity). They also found that PM predicted a significant part of the variance for all temporal fluency measures at T2 (explaining between 4.5% and 9.7% of the L2 fluency variance) and concluded that learners with greater PM capacity achieve higher fluency results, regardless of learning context. However, this study only used one task type to assess PM and no measures of WM to explain the contribution of memory to L2 fluency development.

It is important to note that research on the relationship between WM and PM to L2 fluency has generally been marked by inconsistent findings, many of which may be explained by methodological differences (Juffs & Harrington, 2011). For example, in some studies, tasks like the BDS seem to capture the link between WM and oral production, but little is known about the link between WM capacity and temporal aspects of L2 fluency in a developmental context. PM and L2 fluency may be related; however, the different tests (NWR or SNWR) may capture different constructs (Hummel & French, 2016), making it difficult to determine which aspects of PM (processing and/or storage) are most important to L2 fluency development. Finally, as most studies have only focused on adult populations, the roles of PM and WM in younger learners' fluency remains virtually unexplored, particularly in developmental contexts.

IV Research issues and questions

Research to date has shown a positive relationship between L1 fluency, WM and PM, and L2 fluency, suggesting that features such as speech rate can predict L2 fluency (Towell & Dewaele, 2005). Therefore, we examined the fluency development of adolescent French speakers enrolled in an intensive 9-month English language program. Specifically, we were guided by the following research questions:

- Is there a relationship between L1 and L2 fluency at the beginning (T1) and the end (T2) of an intensive L2 program?

- Is there a relationship between WM (T1) and PM (T1) and L2 fluency at T1 and T2?
- Do memory (WM, PM), L1 fluency and L2 initial fluency level play a significant role in L2 fluency development (T1 to T2), and if so, what proportion of this development do they explain?

V Method

1 Participants

Participants were French-speaking Grade 6 learners (*n* = 47; age: 11) enrolled in an intensive English program (approximately 400 hours of instruction over 9 months) in Quebec, Canada. Background questionnaires showed that all participants attended a francophone public school before the intensive program. None of the students' parents reported speaking English as L1, nor did they report substantial exposure to English outside of the classroom. Participants reported no hearing, speech, or cognitive impairments.

2 Instruments

Speech elicitation tasks: We evaluated learners' L1 and L2 utterance fluency in two tasks by measuring temporal characteristics of speech (identified in the present context as speech rate (SR), phonation time (PT) and mean length of run (MLR)). Task 1 was a picture-cue monologic task based on The Suitcase Story (Derwing et al., 2004), which targets common vocabulary for students in this age group (e.g. physical descriptions, shapes, and colors). Participants had up to two minutes of planning time to prepare what to say but could start when they were ready. We offered no lexical or grammatical prompts during storytelling; however, learners could continue looking at the pictures while narrating the story.

The second task was a semi-structured interview based on the OPI (ACTFL, 2012). We adapted the interview to reflect adolescent experiences (e.g. questions about family and friends, hobbies, summer vacations). Participants answered questions that addressed topics ranging from simple facts (e.g. phone numbers) to more complex subjects (e.g. What would you do if you were a superhero?). Because we used the same tasks in both L1 and L2, we avoided semantic priming by administering speech production tasks first in L2 and then in L1.

3 Memory measures[1]

We assessed WM using the well-documented BDS (Kormos & Sáfár, 2008; Kudo & Lee Swanson, 2014) from the *Wechsler Intelligence Scale for Children* (WISC-IV) (Wechsler, 2003) for children between 6 and 16 years old. The WISC manual reports a Cronbach alpha of .85 for this test indicating high reliability. Participants had to repeat series of an increasing number of items (from 2 to 8) in reverse order.

Table 1. Non-word repetition different structures.

Consonant–vowel structure	Consonant–consonant–vowel structure
pémeudunkinvo	blinpsongliflatran
bofunnangonti	vléplouscungleudro

Note. Total number of consonant–vowel structures = 105 syllables; total number of consonant-consonant–vowel structures = 60 syllables.

Table 2. Example of serial non-word recognition series.

Series	Expected answers
fluie clond **zien cro** prard na ti	different
fluie clond **cro zien** prard na ti	
clat bron prar roir ti	same
clat bron prar roir ti	

Note. Total number of series = 16.

Our first PM measure was a NWR test (Poncelet & Van der Linden, 2003). Following a 4-item practice, the test comprised a total of 28 non-words in series varying from 2 to 8 syllables using two types of syllable structures (see Table 1). We used prerecorded items[2] to eliminate guessing based on lip-reading (Fortier, Simard, & French, 2012).

Our second PM measure was a SNWR task (Abdallah, 2010). Participants listened to 16 pairs of non-word series, varying in length from 5–8 words) and compared them. Learners decided whether the two series in a pair were the same or different; we limited dissimilarity to inverting two words within a given pair of series (see Table 2). Two native French speakers rated the full set of non-word series on a 1–5 scale to measure wordlikeness (1 = ‘very unwordlike’ and 5 = ‘very wordlike’). Wordlikeness was rated low (m = 2.3). After a practice session with the examiner, the test was administered in increasing order (5 to 8 words), four pairs for each length. The test was recorded for subsequent analysis.

4 Procedures

We collected data during the first (T1) and last (T2) weeks of the 9-month intensive English program. We spread the tests over several days to avoid participant fatigue (see Table 3).

We analysed all speech samples using Praat (De Jong & Wempe, 2009). Two-person teams of research assistants transcribed all samples, including repetitions, repairs, fillers, L1 words and inaudible words. The data in this study includes all speech production longer than 2 ms. Three judges manually calculated temporal measures of utterance fluency (SR, PT and MLR) using the audio file and transcription and then categorized each segment as speech, pause, or inaudible. To ensure an optimal threshold for correlation between L2 competency and silent pauses (Bosker et al., 2013), judges noted pause

Table 3. Data collection schedule.

	Time 1				Time 2		
	Day 1	Day 2	Day 3	Day 4	Day 1	Day 2	Day 3
Narration (L1 fluency)			✗				
Interview (L1 fluency)				✗			
Narration (L2 fluency)	✗				✗		
Interview (L2 fluency)		✗				✗	
Backward digit span				✗			✗
Non-word repetition				✗			✗
Serial non-word recognition			✗				✗

durations of up to 250 ms in both L1 and L2. In the modified OPI (ACTFL, 2012), T1 production was minimal for most participants due to low proficiency level. Only two questions elicited production of 10 or more words, and we used these to calculate speech production (see Munro & Derwing, 1999).

Judges scored the BDS test for accuracy, awarding one point per correct answer (max 14). We calculated NWR results for both word and syllable accuracy to facilitate more precise results (Poncelet & Van der Linden, 2003). Two judges scored the NWR test and reached interrater agreement of 93.4% (total words) and 95.2% (total syllables). Due to a technical malfunction, we do not report NWR data from T2. We include the T1 data in our results because they serve as a predictive measure of fluency and allow us to better answer research question 3.

VI Results

Table 4 summarizes L1 temporal measures for both tasks at T1 and reveals a broad range of results among participants, suggesting potential L1 effects. Paired sample *t*-tests comparing task type revealed that both SR ($p < 0.01$) and PT ($p < 0.01$) performance were significantly higher in the interview than the narration task, though there was no significant difference for MLR.

Table 5 compares L2 temporal measures for both tasks at T1 and T2. Results showed a significant improvement for all temporal measures ($p < 0.01$) between T1 and T2 on both tasks, with a large effect size (Cohen's $d = .9$ and above). Like the L1 measures (Table 4), L2 performance also appears to vary as a function of task type. Paired sample t-tests revealed significantly lower scores for SR ($p < 0.01$) and PT ($p < 0.01$) on the narrative task.

Table 6 summarizes results for all WM and PM tests at T1 and T2. The results show that performance on both the BDS and SNWR tasks was statistically similar at both times ($p > .05$), suggesting that WM and PM capacity remained unchanged throughout the study.

To assess the relationship between L1 and L2 fluency at the beginning and at the end of the intensive L2 program, we calculated Spearman correlations for all memory and

Table 4. Descriptive statistics for first language (L1) temporal fluency measures (SR, PT, MLR) narrative and interview tasks at T1.

Measures	Median[a]	Range
Narration (L1):		
SR	195.70	74.50–324.25
PT	71.40	44.89–91.02
MLR	8.14	3.95–20.33
Interview (L1):		
SR	220.02	129.47–278.67
PT	83.69	64.86–97.64
MLR	8.36	3.44–15.17
Composite[b] *(L1)*:		
SR	197.09	146.46–298.44
PT	77.61	63.20–91.17
MLR	8.35	3.70–14.40

Notes. Participants: $n = 47$. * $p < 0.05$. [a] Median reported to mitigate the effects of skewed distribution (see Conover, 1999). [b] Composite score is the average of both tasks (narration and interview). SR = speech rate. PT = phonation time. MLR = mean length of run.

Table 5. Descriptive statistics, *t*-tests, and effect sizes for second language (L2) temporal fluency measures at T1 and T2.

	T1		T2				
Measures	Med	Range	Med	Range	*t*	*p*	*d*
Narration (L2):							
SR	50.49	6.67–98.55	87.25	34.71–144.25	−12.930	0.000*	0.99
PT	28.01	3.98–60.54	49.14	23.48–74.29	−11.631	0.000*	0.98
MLR	2.53	1.30–5.00	2.98	2.00–5.25	−4.259	0.000*	0.92
Interview (L2):							
SR	100.31	38.48–175.79	120.44	64.71–231.74	−6.902	0.000*	0.96
PT	60.70	25.34–92.59	71.46	42.04–90.51	−5.027	0.000*	0.94
MLR	3.02	1.62–7.93	4.18	2.44–11.75	−5.683	0.000*	0.94
Composite[5] *(L2)*:							
SR	72.54	32.63–110.15	108.33	66.08–171.54	−14.688	0.000*	0.99
PT	45.14	20.58–73.28	61.06	38.79–75.10	−11.196	0.000*	0.98
MLR	2.79	1.62–5.54	3.73	2.50–7.09	−6.982	0.000*	0.96

Notes. Participants: $n = 47$. * $p < 0.05$. Composite score is the average of both tasks (narration and interview). SR = speech rate. PT = phonation time. MLR = mean length of run.

fluency measures at T1 and T2. Tables 7 and 8 present results for T1 and T2, respectively.

There was a significant correlation between SR in the L1 OPI interview and the L2 composite score ($\rho = .33, p < 0.01$) at T1. SR in the L1 also correlated significantly with SR in the T2 OPI interview ($\rho = .43, p < 0.01$) and with all temporal measures in the T2

Table 6. Descriptive statistics of first language (L1) working memory (WM) at T1 and T2.

	T1		T2				
Measures	Med	Range	Med	Range	*t*	*p*	*d*
BDS	5	2–8	5	2–8	−0.22	0.83	0.12
SNWR	12	7–15	11	5–15	0.95	0.35	0.42
NWR (words)	12	3–23					
NWR (syllables)	115	61–145					

Note. Participants T1 ($n = 47$). Participants T2 ($n = 37$). BDS maximum score = 14, SNWR maximum score = 16. NWR (words) maximum score = 36. NWR (syllables) maximum score = 165. Cronbach alpha .85, .75, .79, and .77, respectively. * $p < 0.05$.

composite score (SR: $\rho = .39$, $p < 0.01$; PT: $\rho = .30$, $p < 0.05$; MLR: $\rho = .36$, $p < 0.05$). These findings show a clear correlation between L1 and L2 SRs. In the present context, the L1 speed of delivery (interview) was a salient variable in understanding the role of L1 in L2 fluency development.

As for the relationship between WM and PM at T1 and L2 fluency at T1 and at T2, the main results showed that WM (BDS task) was correlated significantly with all temporal measures in the L2 (except for SR in the interview) (narration SR: $\rho = .29$, $p < 0.05$; PT: $\rho = .31$, $p < 0.05$; MLR: $\rho = .38$, $p < 0.01$; interview PT: $\rho = .33$, $p < 0.05$; MLR: $\rho = .31$, $p < 0.05$) and with the composite measure (SR: $\rho = .37$, $p < 001$); PT: $\rho = .40$, $p < 0.01$; MLR: $\rho = .42$, $p < 0.01$). At T2, although the correlations were smaller in magnitude, WM (BDS task) was still significantly correlated with the majority of L2 temporal measures in both tasks, suggesting that the capacity to store and process information simultaneously in memory is related to specific temporal dimensions of L2 fluency.

Further, PM, as referenced by NWR (total syllables), was significantly correlated with MLR in the composite measure both in the L1 ($\rho = .31$, $p < 0.05$) and in the L2 ($\rho = .32$, $p < 0.05$). At T2, NWR (total syllables) was correlated with MLR on both tasks (narration: $\rho = .30$, $p < 0.05$; interview: $\rho = .29$, $p < 0.05$) and with the composite score ($\rho = .37$, $p < 0.01$). These correlations revealed a possible link between PM capacity and MLR in both L1 and L2.

To examine the extent to which memory (WM, PM) and L1 fluency play a role in L2 fluency development (T1 to T2), we conducted standard regression analyses using nine response variables (SR, PT and MLR in the L2 for both tasks and the composite score). For each analysis, the regressors were 1) the control variable in the L2, 2) the three temporal measures in the L1 for the same task, and 3) all memory tests scores (BDS, NWR (syllables) and SNWR). We inserted all variables simultaneously to calculate each variable's relative importance while controlling for all others. We used the Shapiro–Wilk Test (Shapiro & Wilk, 1965) to test the assumption of normality and tested homogeneity of variance using the White Test (1980).

The SR model for the interview (Table 9) reached significance with L1 and L2 temporal measures, explaining a significant percentage of L2 development as measured by

Table 7. Spearman's correlations between temporal fluency measures (SR, PT and MLR) at T1 in the first language (L1) and in the second language (L2) and memory (BDS, NWR and SNWR) at T1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. NaL1-SR	–																					
2. NaL1-PT	.80**	–																				
3. NaL1-MLR	.79**	.83**	–																			
4. InL1-SR	.28	.29*	.26	–																		
5. InL1-PT	.37*	.39**	.31*	.60**	–																	
6. InL1-MLR	.31*	.26	.34*	.69**	.58**	–																
7. CompL1-SR	.88**	.71**	.70**	.67**	.55**	.55**	–															
8. CompL1-PT	.73**	.88**	.73**	.49**	.74**	.44**	.75**	–														
9. CompL1-MLR	.69**	.67**	.80**	.58**	.52**	.79**	.78**	.71**	–													
10. NaL2-SR	.16	.16	.20	.19	.04	−.05	.24	.13	.07	–												
11. NaL2-PT	.19	.28	.27	.09	.03	−.07	.21	.21	.11	.93**	–											
12. NaL2-MLR	.18	.14	.25	.31*	.11	.02	.27	.15	.14	.63**	.50**	–										
13. InL2-SR	−.00	.00	−.05	.27	.36*	.17	.09	.16	.07	.16	.13	.23	–									
14. InL2-PT	.08	.17	.11	.21	.31*	.04	.14	.27	.09	.21	.27	.22	.76**	–								
15. InL2-MLR	.10	.23	.18	.30*	.45**	.22	.17	.38**	.22	.12	.13	.39**	.72**	.65**	–							
16. CompL2-SR	.08	.07	.07	.33*	.28	.13	.21	.18	.12	.66**	.60**	.55**	.80**	.69**	.63**	–						
17. CompL2-PT	.07	.19	.14	.13	.18	−.03	.12	.21	.05	.67**	.74**	.41**	.58**	.80**	.49**	.82**	–					
18. CompL2-MLR	.16	.22	.26	.33*	.34*	.14	.24	.31*	.22	.39**	.34*	.75**	.62**	.57**	.86**	.71**	.57**	–				
19. BDS-T1	−.03	.06	.10	.16	−.00	.04	.01	.04	.09	.29*	.31*	.38**	.19	.33*	.31*	.37**	.40**	.42**	–			
20. SNWR-T1	−.08	−.01	.03	−.11	.09	−.03	−.16	.02	−.03	−.36*	−.32*	−.07	.02	.08	.15	−.14	−.20	.02	.13	–		
21. NWR-Mots-T1	.09	.10	.22	.16	.09	.09	.13	.08	.19	.00	.03	.30*	−.00	.15	.18	.03	.08	.27	.43**	.30*	–	
22. NWR-Syll.-T1	.07	.10	.31*	.15	.02	.21	.08	.06	.31*	−.01	.01	.32*	−.01	.11	.25	.05	.05	.32*	.60**	.37**	.84**	–

Notes. Participants ($n = 47$). * $p < .005$. ** $p < 0.01$. SR = speech rate. PT = phonation time. MLR = mean length of run.

Table 8. Spearman's correlations between temporal fluency measures (SR, PT and MLR) at T2 in the first language (L1) and in the second language (L2) and memory (BDS, NWR and SNWR) at T1.

	NaL1 SR T1	NaL1PT T1	NaL1MLR T1	InL1 SR T1	InL1TP T1	InL1 MLR T1	CompL1SR T1	CompL1PT T1	CompL1MLR T1	1	2	3	4	5	6	7	8	9
1. NaL2-SR_T2	.11	.10	.05	.24	.04	.05	.20	.06	.04	–								
2. NaL2_PT_T2	.22	.27	.21	.20	.07	−.05	.25	.21	.10	.87**	–							
3. NaL2_MLR_T2	.22	.20	.29*	.34*	.09	.28	.33*	.16	.37*	.64**	.59**	–						
4. InL2_SR_T2	−.07	−.05	.00	.43**	.07	.17	.13	.01	.09	.37*	.34*	.25	–					
5. InL2_PT_T2	−.05	.01	.00	.27	.10	.14	.08	.08	.06	.16	.28	.11	.78**	–				
6. InL2_MLR_T2	−.03	.04	.11	.26	.08	.06	.06	.05	.10	.13	.21	.15	.78**	.69**	–			
7. CompL2_SR_T2	.00	.03	.02	.39**	.04	.06	.18	.06	.07	.82**	.73**	.56**	.81**	.56**	.55**	–		
8. CompL2_PT_T2	.17	.26	.22	.30*	.11	.04	.26	.25	.16	.69**	.85**	.51**	.65**	.71**	.53**	.82**	–	
9. CompL2_MLR_T2	.07	.18	.28	.36*	.19	.20	.19	.19	.29*	.36*	.40**	.57**	.69**	.54**	.84**	.65**	.60**	–
10. BDS_T1										.26	.32*	.24	.33*	.13	.34*	.33*	.33*	.36*
11. SNWR_T1										−.11	.00	−.01	−.00	.19	.07	−.08	.07	.02
12. NWRmots_T1										−.03	.11	.15	.11	.09	.16	.04	.17	.16
13. NWRSyll_T1			*							−.02	.12	.30*	.16	.13	.29*	.07	.20	.37**

Notes. Participants ($n = 47$). * $p < 0.05$. ** $p < 0.01$. SR = speech rate. PT = phonation time. MLR = mean length of run. BDS = backward digit span. SNWR = serial non-word recognition. NWR = non-word repetition.

Table 9. Detailed model of regression analysis on SR (dependent variable) in the interview at T2.

Model		Unstandardized coefficient		Standardized coefficient	*t*	*p*	*lmg*
	df	B	Standard error	ß			
SR in the second language (L2) at T1	1	0.563	0.150	0.477	3.77	0.001	0.230
SR in the first language (L1)	1	0.530	0.160	0.589	3.30	0.002	0.151
PT in the L1	1	−1.809	0.712	−0.397	−2.54	0.015	0.042
MLR in the L1	1	−1.122	1.905	−0.100	−0.59	0.559	0.017
NWR (syllable) T1	1	0.094	0.273	0.057	0.34	0.733	0.012
BDS T1	1	1.898	3.259	0.089	0.58	0.564	0.060
SNWR T1	1	0.493	2.116	0.031	0.23	0.817	0.002

Notes. Participants ($n = 47$). (F (7.39) = 5.90, p = 0.000 (r2 = 0.514), W = 0.977788, p = 0.5042).
*$p < 0.05$. SR = speech rate. PT = phonation time. MLR = mean length of run. NWR = non-word repetition. BDS = backward digit span. SNWR = serial non-word recognition.

SR. This indicates that L1 speed of delivery can predict the later L2 speed of delivery. Other regression analyses showed that all temporal measures of L2 fluency at T1 (except for MLR in the narration) were significant predictors of all temporal measures of L2 fluency development between T1 and T2 (explaining between 14.5% to 34.8% of the variation). In the L1, only SR in the interview reached significance, explaining 15.1% of the SR variance in the interview between T1 and T2, while none of the memory variables reached significance.

VII Discussion

In this study, we examined the contribution of L1 fluency, WM, PM, and initial L2 fluency to subsequent L2 fluency development during an intensive L2 English program. We evaluated L1 and L2 utterance fluency using temporal measures (SR, PT, and MLR) in two task types (narrative and interview) at two points in time over a 9-month period. Overall, the study revealed that L1 fluency, WM, PM, and initial L2 fluency contributed to L2 fluency development, that L2 temporal measures and L1 SR were significant predictors of L2 fluency development, and that results varied as a function of task type.

Our findings revealed that L2 fluency improved considerably (with effect sizes greater than .9) between T1 and T2, thus adding to the overall picture of the impact of intensive English contexts on learning, specifically the development of oral skills (Collins & White, 2011, 2012; Lightbown & Spada, 1994; White & Turner, 2005). Furthermore, it is important to note that these results also provide support to the growing body of evidence that initial L2 fluency is an important predictor of L2 fluency development as reported in previous studies that used similar measures in similar contexts (e.g. study abroad) with comparable effect sizes. Finally, results suggested that the temporal measures used in the current study (SR, PT, and MLR) were sensitive to improvement at the early stages of L2 fluency development.

1 The relationship between L1 and L2 fluency

Our first research question examined the specific relationship between L1 and L2 fluency at T1 and 9 months later at T2. In the narration task, the strength of the correlation between L1 and L2 MLR decreased between T1 and T2, while in the interview the strength of the correlation between L1 and L2 SR increased between T1 and T2. Towell and Dewaele (2005) also found a relationship between SR in the L1 and SR in an L2 narration task. One plausible explanation for this apparent task effect is that more advanced learners (such as the population in Towell & Dewaele) may align their production in L2 with their L1 in different tasks. In contrast, beginners are unable to differentiate performance given their lower L2 proficiency level.

While we did not find a significant correlation between L1 SR and L2 SR in the narration task, our results were comparable to Derwing et al. (2009), who found decreasing correlations between L1 temporal measures (pauses per second, SR, and pruned syllables per second) and the same L2 measures over time. Overall, our results are consistent with a growing body of research highlighting performance differences between monologic and dialogic tasks (e.g. Préfontaine & Kormos, 2015; Skehan et al., 2016).

Our findings of a task effect also support the hypothesis that conflicting results in previous research may be the result of specific tasks (Kormos & Trebits, 2012; Préfontaine & Kormos, 2015). In monologic tasks (e.g. narration), speakers rely on their own knowledge to formulate speech. This task type may increase the processing burden and restrict lower proficiency learners' access to essential language components (e.g. restricted vocabulary and logical order of events in a picture-cue narrative), potentially masking the relationship between L1 and L2. Dialogic tasks (e.g. interview), decrease the processing load most likely because speakers can rely on their own knowledge and their interlocutor's speech to conceptualize a message, which may lead to greater cognitive fluency and to L2 production that shares more features with their L1. Future research could examine the role of cognitive fluency and the relationship between L1 and L2 fluency while avoiding the interlocutor effect by employing monologic tasks that elicit greater speech production (e.g. story recall).

2 The relationship between memory and L2 fluency

The second research question addressed the relationship between WM, PM and L2 fluency. The results showed moderate correlations (ranging from .29 to .42) at T1 and T2 between WM capacity (BDS) and all temporal measures of L2 fluency (except for SR in the interview); there was also a link between PM capacity (NWR (total syllables)) and MLR in both tasks. Our results were consistent with previous findings in which WM was related to the capacity to form longer units from chunks (Zhang & Simon, 1985), to SR (Neilson, 2014), and to speech production (Kormos & Sáfár, 2008). They were also consistent with previous L1 findings that performance in digit span tests was related to more spontaneous and higher quality utterances (Adams & Gathercole, 1995, 2000). Our findings suggest that increases in WM capacity are related to utterance length and overall speaking time. In terms of PM capacity, current findings showed a moderate relationship between NWR (total syllables) and MLR in the monologic task at T1 and T2 and in the interview task at T2. Our results support O'Brien et al. (2007) who found a correlation between PM capacity (using a SNWR task) and the longest run without pauses at T2.

However, our results are inconsistent with Kormos and Sáfár's (2008) finding of no relationship between participants' NWR scores and L2 fluency. This difference may be explained by Kormos and Sáfár's use of perceived fluency measures (using the criteria of the Cambridge First Certificate oral exam rubric) rather than temporal fluency measures. Furthermore, prior to testing (T1), Kormos and Sáfár's participants had received more language instruction than those in the present study and may have reached a level at which PM contribution started to decrease. As a result, although our participants' fluency level increased significantly, it remained within a range where PM contribution was still significant. However, the weak correlation between PM and L2 fluency suggests that the contribution of PM had already begun to decrease for these learners in this context.

Methodological differences may explain the often contradictory results in this line of research (Juffs & Harrington, 2011). Some studies have investigated the role of PM using measures in L1 (Kormos & Sáfár, 2008) or L2 (Hummel & French, 2016), using different tasks such as NWR (e.g. Kormos & Sáfár, 2008) or SNWR (e.g. O'Brien et al., 2007). Researchers have generally used either monologic (Derwing et al., 2004, 2009; Préfontaine & Kormos, 2015) or dialogic tasks (O'Brien et al., 2007). The use of different temporal measures used to examine the relationship between fluency and other variables affects the magnitude of correlations (e.g. Huensch & Tracy-Ventura, 2017; Segalowitz, French, Guay, & Callahan, 2014). Future research would benefit from an exploration of the validity of various temporal predictors of L2 fluency.

3 Predictors of L2 fluency development

The final research question concerned the relative contribution of L1 fluency, WM, and PM to L2 fluency development. The main findings showed that SR in the L1 interview and all L2 measures at T1 (except narration MLR) were significant predictors of L2 fluency. These findings suggest that learners' initial L2 fluency level was a more important predictor of later L2 fluency development than other variables in this study. A potential explanation for this finding is that learners with higher levels of fluency developed prior to the intensive language program have an advantage in the proceduralization of their existing ability (i.e. skill acquisition theory; Dekeyser, 2007, 2015; Suzuki & Dekeyser, 2017). Intensive learning contexts allow frequent opportunities for intensive L2 practice in a variety of communicative situations (e.g. Collins et al., 1999; Collins & White, 2011, 2012; Lightbown, 2001, 2003; Lightbown & Spada, 1991, 1994, 1997; Spada & Lightbown, 1989). French, Gagné and Collins (2020) also noted that the fluency measures used in this study are especially sensitive to proceduralization effects (Goldman Eisler, 1968). Consequently, it is possible that our study captured the proceduralization of declarative knowledge (e.g. words and chunks) gained in formal classroom settings prior to the intensive program (i.e. T1).

Regarding the role of L1 fluency in L2 fluency development, SR in the L1 interview was a significant predictor of SR in the L2 interview. SR contributed significantly to the model, explaining 15.1% of variation in L2 fluency development. This result is consistent with previous findings regarding the influence of L1 fluency on L2 fluency (De Jong et al., 2015; Huensch & Tracy-Ventura, 2017). One potential explanation for this finding may be Cummins' (2001; Cummins & Swain, 1986) iceberg theory on the interdependence of languages, where languages share cognitive processes, including those underlying oral fluency. This could explain why the SR capacity in one language could predict SR in another language.

Our findings regarding WM were inconclusive. Few studies have examined the role of WM in L2 fluency and those that have examined it have reported inconsistent results. For example, Mackey et al. (2010) found that WM explained a portion of speech production while Mizera (2006) found no link between WM and L2 production. These inconsistencies indicate a need for additional investigation of the role of WM in L2 fluency development.

Finally, our results regarding the role of PM in L2 fluency development are inconsistent with those reported by O'Brien et al. (2007). These authors found that PM predicted up to 9.7% of the variation in L2 fluency. However, because our methodologies differ, a comparison with this study was not feasible.

VIII Implications and conclusions

This study highlights significant differences in students' performance on different task types and provides further support to previous research that has examined the contribution(s) of L1 fluency, WM, PM, and initial L2 fluency to subsequent L2 fluency development. We extended this line of inquiry to the intensive classroom learning context. The current findings are important for language instructors, especially when they assess their students' performance. Monologic tasks increase processing load and adversely impact L2 fluency output, which may lead to an incomplete picture of learner fluency. On the other hand, dialogic tasks decrease processing load and positively impact L2 fluency output, which may lead to an overestimation of learner fluency. Therefore, instructors should carefully consider task effects on fluency when developing assessment measures.

We also found that L1 SR in a dialogic task was a significant predictor of later L2 fluency development, while L1 SR in a monologic task was not. This finding indicates that eliciting speech using more than one task type may provide a more complete picture of the relationship between variables (e.g. L1 fluency, WM, and PM). This study also highlights the importance of initial L2 fluency as a predictor of later L2 fluency development. Future research could extend the assessment of L2 fluency as a baseline and investigate the nature of L2 speech production to determine if other individual differences help explain L2 fluency development.

The present study is not without limitations. Our finding that SNWR did not correlate with temporal measures should be addressed in future research. While there may ultimately be a relationship between SNWR and fluency, the homogeneity of SNWR scores in our population may have prevented us from detecting this potential relationship. A second limitation of our study is that we used the same task to assess L1 and L2 fluency.[3] We did this to facilitate comparison across performance in both languages at T1 and T2. Although nine months elapsed between tests, it is possible that there was a potential practice effect.

Nevertheless, despite its limitations, this study is the first to our knowledge to have investigated the relation between L1 and L2 fluency using different tasks in an intensive classroom setting. The main findings point to the important role of L1 individual differences, such as SR, in fluency development and suggest that SR, PT and MLR may be particularly reliable measures of L2 fluency development at early stages of acquisition. Moreover, fluency development was not only explained by L1 fluency and differences in WM and PM, but also by L2 fluency levels. Finally, L2 fluency development varied as a function of task type, with monologic tasks placing the heaviest processing load on temporal measures of utterance fluency. Therefore, it may be beneficial for future research to further investigate the nature of initial L2 fluency levels by examining overall L2 proficiency in developmental contexts, using different task types, to shed light on other factors that may predict L2 development at different stages of acquisition.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Nancy Gagné https://orcid.org/0000-0003-0996-2540

Notes

1. Because previous research indicates that testing WM in L2 can potentially be a confound with L2 expertise (Linck et al., 2014), we administered all assessments of WM and PM in L1. Studies have argued that increasing L2 knowledge, along with WM scores in the L2, would produce an overlapping measure of both variables (Cheung, 1996; Kormos & Sáfár, 2008; Linck et al., 2014; Masoura & Gathercole, 1999), making memory contributions difficult to interpret.
2. All study recordings occurred in an anechoic chamber at a frequency of 44,000 Hz using Audacity (2012).
3. Analysis of the scatter plots showed that some scores were higher in the L2 at T2 than in the L1 at T1. The distribution revealed that none of the participants had higher scores in the L2 for all temporal measures and that these scores were attributed to different participants. A test effect may explain this finding, as the same tasks were used at both T1 and T2; however, it is likely that participants were more mature and at ease (more confident) at T2, leading to greater speech production in the L2. This finding highlights the importance for future research to assess the L1 at T2 in order to determine potential changes in L1 temporal measures between T1 and T2.

References

Abdallah, F. (2010). *The role of phonological memory in L2 acquisition in adults at different proficiency levels*. Unpublished doctoral dissertation, Université Laval, Québec, QC G1V 0A6, Canada.

ACTFL. (2012). *Oral Proficiency Interview (OPI)*: Language Testing International.

Adams, A.M., & Gathercole, S. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research*, *38*, 403–414.

Adams, A.M., & Gathercole, S. (2000). Limitations in working memory: Implications for language development. *International Journal of Language and Communication Disorders*, *35*, 95–116.

Audacity, T. (2012). *Audacity(R):Free Audio Editor and Recorder* (Version 2.0.0) [Computer application]. https://audacityteam.org.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*, 189–208.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, A., & Hitch, G. (1974). Working memory. In Bower, G.H. (Ed.), *The psychology of learning and motivation: Advances in research and theory: Volume 8* (pp. 47–89). Academic Press.

Baddeley, A., Gathercole, S.E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173.

Birdsong, D. (2005). Interpreting age effects in second language acquisition. In Kroll, J.F., & A.M.B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic perspectives* (pp. 109–127). Oxford University Press.

Birdsong, D. (2006). Dominance, proficiency, and second language grammatical processing. *Applied Psycholinguistics*, *27*, 46–49.

Birdsong, D. (2009). Age and the end state of second language acquisition. In Bhatia, W.R.T. (Ed.), *The new handbook of second language acquisition*. Elsevier.

Birdsong, D. (2014). Dominance and age in bilingualism. *Applied Linguistics*, *35*, 374–392.

Bosker, H.R., Pinget, A.-F., Quené, H., Sanders, T., & de Jong, N.H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, *30*, 159–175.

Breiner-Sanders, K., Lowe, P., Miles, J., & Swender, E. (2000). ACTFL proficiency guidelines: Speaking, revised 1999. *Foreign Language Annals*, *33*, 13–18.

Brooks, P.J., Kempe, V., & Sionov, A. (2006). The role of learner and input variables in learning inflectional morphology *Applied Psycholinguistics*, *27*, 185–209.

Chambers, F. (1997). What do we mean by fluency? *System*, *25*, 535–544.

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology*, *3*, 867–873.

Collins, L., & White, J. (2011). An intensive look at intensity and language learning. *TESOL Quarterly*, *45*, 107–133.

Collins, L., & White, J. (2012). Closing the gap: Intensity and proficiency. In Munoz, C. (Ed.), *Intensive exposure experiences in second language learning* (pp. 45–65). Multilingual Matters.

Collins, L., Lightbown, P.M., Halter, R.H., & Spada, N.M. (1999). Time and the distribution of time in L2 instruction. *TESOL Quarterly*, *33*, 655–680.

Conover, W.J. (1999). *Practical nonparametric statistics*. 3rd edition. John Wiley & Sons Inc.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford University Press.

Cucchiarini, C., Strik, H., & Boves, L. (2000). Quantitative assessment of second language learners' fluency by means of automatic speech recognition technology. *Journal of the Acoustical Society of America*, *107*, 989–999.

Cucchiarini, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: Comparisons between reading and spontaneous speech. *Journal of the Acoustical Society of America*, *111*, 2862–2873.

Cummins, J. (2001). The entry and exit fallacy in bilingual education. In C. Baker & N. H. Hornberger (Eds.), *An introductory Reader to the Writings of Jim Cummins* (pp. 110–138). Multilingual Matters.

Cummins, J., & Swain, M. (1986). *Bilingualism in Education*. Longman Press.

De Jong, N.H., & Mora, J.C. (2017). Does having good articulatory skills lead to more fluent speech in first and second languages? *Studies in Second Language Acquisition*, *41*, 227–239.

De Jong, N.H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*, 385–390.

De Jong, N.H., Groenhout, R., Schoonen, R., & Hulstijn, J.H. (2015). Second language fluency: speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, *36*, 223–243.

Dekeyser, R. (2007). Study abroad as foreign language practice. In Dekeyser, R. (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

Dekeyser, R. (2015). Skill acquisition theory. In VanPatten, B., & William, J. N. (Eds.), *Theories in second language acquisition*. *An introduction* (pp. 94–112). Routledge.

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., Munro, M.J., Thomson, R.I., & Rossiter, M.J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, *31*, 533–557.

Derwing, T.M., Rossiter, M.J., Munro, J.M., & Thomson, R.I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, *54*, 655–679.

Duran-Karaoz, Z., & Tavakoli, P. (2020). Predicting L2 fluency from L1 fluency behavior: The case of L1 Turkish and L2 English speakers. *Studies in Second Language Acquisition*, *42*, 671–695.

Ellis, N.C. (2001). Memory for language. In Robinson, P. (Ed.), *Cognition and second language instruction* (pp. 33–68). Cambridge University Press.

Ellis, N. C., & Ferreira, J. F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *Modern Language Journal*, 93, 370–385.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, *9*, 147–171.

Fortier, V., Simard, D., & French, L.M. (2012). La mesure de la mémoire phonologique auprès d'une population d'enfants linguistiquement hétérogène [Measuring phonological memory in a linguistically heterogeneous population of children]. *Canadian Modern Language Review / Revue canadienne des langues vivantes*, *68*, 291–315.

Fortkamp, M.B.P. (1998). Measures of working memory capacity and L2 oral fluency. *Ilha do Desterro*, *35*, 201–238.

Fortkamp, M.B.P. (1999). Working memory capacity and aspects of L2 speech production *Communication and Cognition*, *32*, 259–296.

Fortkamp, M.B.P. (2000). Working memory capacity and L2 speech production: An exploratory study. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Santa Catarina, Brazil.

Foster, P. (2020). Oral fluency in a second language: A research agenda for the next ten years. *Language Teaching*, *53*, 446–461.

French, L.M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, *29*, 463–487.

French, L.M., Gagné, N., & Collins, L. (2020). Long-term effects of intensive instruction on fluency, comprehensibility and accentedness. *Journal of Second Language Pronunciation*, *6*, 380–401.

Gathercole, S. E., & Baddeley, A. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200–213.

Gathercole, S.E., & Baddeley, A. (2014). *Working memory and language*. Psychology Press.

Gathercole, S.E., Frankish, C.R., Pickering, J.B., & Peaker, S.M. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 84–95.

Gathercole, S.E., Pickering, S.J., Hall, M., & Peaker, S.M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology: Section A*, *54*, 1–30.

Gathercole, S.E., Tiffany, C., Briscoe, J., Thorn, A.S.C., & ALSPAC Team (2005). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *46*, 598–611.

Gathercole, S.E., Willis, C.S., Emslie, H., & Baddeley, A. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, *12*, 349–367.

Georgiadou, E., & Roehr-Brackin, K. (2017). Investigating executive working memory and phonological short-term memory in relation to fluency and self-repair behavior in L2 speech. *Journal of Psycholinguistic Research*, *46*, 877–895.

Gilabert, R., & Muñoz, C. (2010). Differences in attainment and performance in a foreign language: The role of working memory capacity. *International Journal of English Studies*, *10*, 19–42.

Goldman Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. University of Michigan.

Hilton, H. (2008a). Connaissances, procédures et production orale en L2 [Knowledge, procedures and oral production in L2]. *Acquisition et interaction en langue étrangère*, *27*, 63–89.

Hilton, H. (2008b). The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal*, *36*, 153–166.

Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance et and proficiency: complexity, accuracy and fluency in SLA.* Amsterdam: John Benjamins Publishing Company.

Huensch, A., & Tracy-Ventura, N. (2017). Understanding second language fluency behavior: The effects of individual differences in first language fluency, cross-linguistic differences, and proficiency over time. *Applied Psycholinguistics*, *38*, 755–785.

Hulme, C., Maughan, S., & Brown, J.D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*, 685–701.

Hulstijn, J.H., & Bossers, B. (1992). Individual differences in L2 proficiency as a function of L1 proficiency. *European Journal of Cognitive Psychology*, *4*, 341–353.

Hummel, K.M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics*, *30*, 225–249.

Hummel, K.M., & French, L.M. (2016). Phonological memory and aptitude components: Contributions to second language proficiency. *Learning and Individual Differences*, *51*, 249–255.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, *44*, 137–166.

Kieras, D.E., Meyer, D.E., Mueller, S., & Seymour, T. (1999). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual-motor and cognitive human performance. In Miyake, A., & P. Shah (Eds.), *Models of working memory mechanisms of active maintenance and executive control* (pp. 183–223). Cambridge University Press.

Kormos, J. (2006). *Speech production and second language acquisition*. Routledge.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, *11*, 261–271.

Kormos, J., & Trebits, A. (2012). The role of task complexity, modality and aptitude in narrative task performance. *Language Learning*, *62*, 439–472.

Kudo, M., & Lee Swanson, H. (2014). Are there advantages for additive bilinguals in working memory tasks? *Learning and Individual Differences*, *35*, 96–102.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40*, 387–417.

Lightbown, P.M. (2001). L2 instruction: Time to teach. *TESOL Quarterly*, *35*, 598–599.

Lightbown, P.M. (2003). SLA research in the classroom/SLA research for the classroom. *Language Learning Journal*, *28*, 4–13.

Lightbown, P.M., & Spada, N.M. (1991). Étude des effets à long terme de l'apprentissage intensif de l'anglais, langue seconde, au primaire [Study of the long-term effects of intensive learning of English as a second language in elementary school]. *Canadian Modern Language Review / Revue canadienne des langues vivantes*, *48*, 90–117.

Lightbown, P.M., & Spada, N.M. (1994). An innovative program for primary ESL students in Quebec. *TESOL Quarterly*, *28*, 563–579.

Lightbown, P.M., & Spada, N.M. (1997). Learning English as a second language in a special school in Quebec. *Canadian Modern Language Review*, *53*, 315–355.

Linck, J.A., Osthus, P., Koeth, J.T., & Bunting, M.F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin and Review*, *21*, 861–883.

Lord, G. (2006). Defining the indefinable: Study abroad and phonological memory abilities. In Klee, C.A., & T.L. Face (Eds.) *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 40–46). *Cascadilla Proceedings Project*.

Mackey, A., R., Adams, C., Stafford, P., & Winke. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, *60*(3), 501–533.

Marini, A., Eliseeva, N., & Fabbro, F. (2019). Impact of early second-language acquisition on the development of first language and verbal short-term and working memory. *International Journal of Bilingual Education and Bilingualism*, *22*(2), 165–176. doi: 10.1080/13670050.2016.1238865

Masoura, V.M., & Gathercole, S.E. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology*, *34*, 383–388.

Mizera, G.J. (2006). Working memory and L2 oral fluency. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA, USA.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, *44*(1), 1–35. doi: 10.1017/S0261444810000327

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*, 285–310.

Munro, M.J., & Derwing, T.M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*, 285–310.

Neilson, K.B. (2014). Can planning time compensate for individual differences in working memory capacity? *Language testing Research*, *18*, 272–293.

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, *27*, 377–402.

O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, *29*, 557–581.

Peltonen, P. (2018). Exploring connections between first and second language fluency: A mixed methods approach. *The Modern Language Journal*, *102*, 676–692.

Poncelet, M., & Van der Linden, M. (2003). L'évaluation du stock phonologique de la mémoire de travail: élaboration d'une épreuve de répétition de non-mots pour population francophone [Evaluation of the phonological stock of working memory: Development of a non-word repetition test for the French-speaking population]. *Revue de neuropsychologie*, *13*, 377–407.

Préfontaine, Y., & Kormos, J. (2015). The relationship between task difficulty and second language fluency in French: A mixed methods approach. *Modern Language Journal*, *99*, 96–112.

Riazantseva, A. (2001). Second language proficiency and pausing : a study of Russian speakers of English. *Studies in Second Language Acquisition*, *23*(04), 497–526. doi: 10.1017/S027226310100403X

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, *14*, 423–441.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, *25*, 46–73.

Segalowitz, N. (2010). *The cognitive bases of second language fluency*. Routledge.

Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *IRAL: International Review of Applied Linguistics in Language Teaching*, *54*, 79–95.

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*(02), 173–199.

Segalowitz, N., French, L.M., Guay, J.D., & Callahan, M. (2014). A la recherche des caractéristiques critiques de l'aisance à l'oral en français langue seconde [In search of the critical features of oral fluency in French as a second language]. Unpublished paper presented at the 'Les perspectives actuelles sur le développement de la compétence à l'oral en français langue seconde', annual meeting of the Association francophone pour le savoir (ACFAS), Montréal, QC.

Segalowitz, N., & Hulstijn, J. H. (2005). Automaticity in bilingualism and second language learning. In J. Kroll & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). Oxford, UK.: Oxford University Press.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52((3/4)), 591–611.

Skehan, P., Foster, P., & Shum, S. (2016). Ladders and snakes in second language fluency. *IRAL: International Review of Applied Linguistics in Language Teaching*, *54*, 97–111.

Spada, N., & Lightbown, P.M. (1989). Intensive ESL programmes in Quebec primary schools. *TESL Canada Journal*, *7*, 11–32.

Speciale, G., Ellis, N.C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics*, *25*, 293–321.

Suzuki, Y., & Dekeyser, R. (2017). The interface of explicit and implicit knowledge in second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, *67*, 747–790.

Towell, R., & Dewaele, J.-M. (2005). The role of psycholinguistic factors in the development of fluency amongst advanced learners of French. In Dewaele, J.-M. (Ed.), *Focus on French as a foreign language: Multidisciplinary approaches* (pp. 210–239). Multilingual Matters.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced of French. *Applied Linguistics*, *17*, 84–119.

Trebits, A., & Kormos, J. (2008). Working memory capacity and narrative task performance. Unpublished paper *presented at the 33rd international LAUD symposium*, Essen, Germany.

Vallar, G., & Baddeley, A.D. (1984a). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behaviour*, *23*, 151–162.

Vallar, G., & Baddeley, A.D. (1984b). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case study. *Cognitive Neuropsychology*, *1*, 121–141.

Valls-Ferrer, M., & Mora, J.C. (2014). L2 fluency development in formal instruction and study abroad. In Pérez-Vidal, C. (Ed.) *Language acquisition in study abroad and formal instruction contexts*, (pp. 111–136). John Benjamins.

Vulchanova, M., Foyn, C.H., Nilsen, R.A., & Sigmundsson, H. (2014). Links between phonological memory, first language competence and second language competence in 10-year-old children. *Learning and Individual Differences*, *35*, 87–95.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children – 4th Edition (WISC-IV)*. Harcourt Assessment.

White, J.L., & Turner, C.E. (2005). Comparing children's oral ability in two ESL programs. *Canadian Modern Language Review*, *61*, 491–517.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*, 817–838.

Zabihi, R., Ghominejad, S., & Javad Ahmadian, M. (2021). Can willingness to communicate, communication in English anxiety, behavioural inhibition and behavioural action predict perceived l2 fluency? *Language Teaching Research*. Epub ahead of print 20 September 2021. DOI: 10.1177/13621688211044071.

Zhang, G., & Simon, H.A. (1985). STM capacity for Chinese words and idioms: Chunking and the acoustic loop hypothesis. *Memory and Cognition*, *13*, 93–201.