Learning French sound/spelling links in English primary school classrooms

ALISON PORTER
University of Southampton

ABSTRACT

EN Two classes (n = 45) of 9- to 11-year-old children in different English primary schools experienced weekly French phonics instruction as part of a foreign language (FL) learning programme. The underpinning rationale of the phonics instruction was that systematic and explicit FL phonics could contribute to the development of FL phonological decoding, operationalised as reading aloud individual word cards. The study explored progression in learning French sound/spelling links over this period with data from a sub-sample of students (n = 23) from both classes. Additional variables such as FL general proficiency and FL literacy, as well as L1 reading age, spelling age, and teacher assessed literacy levels were collected to explore relationships with the development of French phonological decoding. This study found that French sound/spelling links developed slowly but showed significant gains between mid- and post-test and that learning appeared resistant to attrition evidenced by delayed post-test scores. Successful phonological decoders were likely to be successful FL learners generally and proficient in L1 literacy.

Key words: YOUNG FL LEARNERS, EARLY FL LITERACY INSTRUCTION, FL PHONICS

ES Dos clases (n = 45) de niños de 9-11 años en diferentes escuelas primarias de inglés experimentaron semanalmente instrucción fónica francesa como parte de un programa de una lengua extranjera (LE). Este modelo pedagógico se implementó porque la instrucción fonética sistemática y explícita de una LE podría contribuir al desarrollo de la decodificación fonológica en la Le, esta última medida en la lectura en voz alta de tarjetas de una sola palabra. El estudio observó la progresión en el aprendizaje de sonido/ortografía con datos de una submuestra de estudiantes (n = 23) de ambas clases. Otras variables fueron consideradas para observar su correlación con el desarrollo de la decodificación fonológica en francés: el nivel de competencia lingüística y alfabetización en la LE, el nivel de lectura en la primera lengua, el nivel de ortografía y los niveles de alfabetización del profesor. Con este estudio se descubrió que la relación sonido-grafía se desarrolló despacio, pero con resultados positivos entre las pruebas durante y después de la instrucción. Además, el aprendizaje se resistió al fenómeno de atrición, como mostraron los resultados de las segundas pruebas después de la instrucción. Los que decodificaban los aspectos fonológicos con éxito tenían más probabilidades de aprender una LE con éxito y con una competencia avanzada en su lengua materna.

Palabras clave: JÓVENES ESTUDIANTES DE LE, APRENDIZAJE TEMPRANO DE UNA LE, INSTRUCCIÓN FÓNICA EN LA LE

IT Come parte di un programma per l'apprendimento di una lingua straniera (LS), due classi (n = 45) di bambini tra i 9 e gli 11 anni in due scuole primarie inglesi hanno sperimentato settimanalmente lezioni di francese con il metodo fonetico. L'ipotesi alla base del metodo è che l'istruzione fonetica sistematica ed esplicita in LS contribuisca allo sviluppo della decodificazione fonologica in LS espressa dalla lettura ad alta voce di schede con una singola parola. Lo studio analizza la progressione nell'apprendimento della corrispondenza fonema/grafema attraverso i dati di un sotto-campione di studenti (n = 23) ins entrambe le classi. Sono stati raccolti dati aggiuntivi relativi a variabili quali la competenza complessiva e l'alfabetizzazione in LS, insieme all'età di lettura in L1, l'età di scrittura e il livello di alfabetizzazione, stabilito dall'insegnante, al fine di analizzarne il rapporto con lo sviluppo della decodificazione fonologica in francese. Da questo studio emerge che le corrispondenze fonema/grafema in francese si sviluppano lentamente ma evidenziano progressi significativi tra il test in itinere e il post-test e che quanto appreso pare resistere all'attrito come dimostrato dai punteggi del post-test differito. Inoltre, è risultato che apprendenti abili nella decodificazione fonologica avessero un'alta probabilità di essere apprendenti di LS complessivamente competenti e con un alto livello di alfabetizzazione in L1.

Parole chiave: GIOVANI APPRENDENTI DI LS, ALFABETIZZAZIONE IN UNA LS, METODO FONETICO IN UNA LS

✉ **Alison Porter,** University of Southampton
A.Porter@soton.ac.uk

1. Literature review
1.1. The importance of FL literacy in England
Foreign language study is now compulsory in English primary schools and was included in the Primary National Curriculum (DfE, 2013) from Key Stage 2 (ages 7-11). The curriculum aims to develop practical communication skills in one language, to create foundational language skills as preparation for ongoing language learning and to develop cultural and intercultural awareness. It includes FL literacy (reading and writing) and requires children to "discover and develop an appreciation of a range of writing" (p. 193), as well as practical skills such as: understanding and responding to written, authentic text (DfE, 2013). There are specific attainment targets which relate to reading for meaning and reading aloud which include the ability to link sounds, spellings and word meaning.

The importance of FL literacy as a component of the development of practical language skills and cultural awareness is widely accepted (Koutsompou, 2015). However, little is known about the efficacy of particular approaches to teaching literacy in FL school settings such as in the UK. Research has found that reading strategy instruction supported reading comprehension and FL motivation in English secondary schools (Macaro & Erler, 2008) and reading comprehension in English primary schools (Macaro & Mutton, 2009). A recent, large-scale project in secondary schools examined to what extent three different pedagogic approaches—phonics instruction, strategy instruction, or text experience without explicit instruction—could contribute to FL (French) linguistic and non-linguistic outcomes. All approaches noted statistically significant improvements in FL reading aloud, FL vocabulary, and FL reading comprehension (Woore, Graham, Porter, Courtney, & Savory, 2018). However, the phonics group made more progress at FL reading aloud than the other conditions, and the strategies group improved most for self-efficacy in French reading (Woore et al., 2018).

Learning FL sound/spelling links in English schools is noted to develop slowly (Cable et al., 2010; Woore, 2011) and unsystematically (Woore, 2011). French reading aloud was particularly problematic for primary school learners, who tended to be influenced by knowledge and application of L1 sound/spelling links (Cable et al., 2010; Porter, 2014). Similarly, in secondary schools, Woore (2007) found that after 18 phonics lessons only the two strongest L1 readers made successful, unscaffolded analogies to single graphemes. He also noted that the weaker L1 readers tended to use larger syllabic units and attributed this to less well-developed phonemic awareness. In other words, a period of focused FL phonics instruction did not appear to compensate for the influence of L1 knowledge and skills.

As FL reading and writing are important in the development of FL skills and attitudes towards FL learning, it is essential that research continues to investigate the effectiveness of particular pedagogic tools and the process of learning literacy in FL classrooms. Phonics is often practised, as a pedagogic tool, in FL settings based on assumptions grounded in the popularity of phonics instruction in L1 English classrooms and evidence concerning its contribution to word recognition skills in alphabetic languages (Adams, 1990; Castles, Rastle & Nation, 2018). From a process perspective, whilst L2 reading has been researched and theorised with bilingual and multilingual populations, there is little evidence about the processes or outcomes for learning to read in FL settings. Again, assumptions cannot be made that pedagogic tools and learning processes identified in L1 and L2 settings will be readily transferable to beginner FL classrooms, known to lack the rich and regular input which supports learning in L1 and L2 settings (Garton, Copland & Burns, 2011; Graham et al., 2017).

1.2. Learning to read in a first language (L1): the importance of word recognition
It is now widely accepted that rapid and accurate word recognition is a fundamental skill in reading for meaning (Adams, 1990; Oakhill & Beard, 1999; Stanovich & Stanovich, 1999; Vellutino & Fletcher, 2005). Word recognition is accessed by either a) decoding spellings into sounds then retrieving meaning (sub-lexical route) or b) retrieving meaning directly from spellings (lexical route) (Castles, Rastle, & Nation, 2018).

1.2.1. The process of developing L1 word recognition
Phonological decoding, involving the mapping of letters or groups of letters onto corresponding sounds (Ziegler, Perry, & Zorzi, 2014), is essential in word recognition. Sight word recognition is believed to develop in stages or "overlapping waves" of phonological decoding, starting with a partial alphabetic phase where connections are made between a few sounds and their corresponding letters, usually initial and final consonants (Ehri, 2014, p. 5; Ehri, 1999). This development is accompanied by spelling development which

starts with written words generated from a relatively small number of learned letters e.g. giraffe = JRF (Castles et al., 2018). Development of links between sound and print is known as "orthographic mapping" (Ehri, 2014: 5) or "orthographic learning" and comprises both word-specific links between sound and print, along with awareness of the probabilistic occurrence of letters and letter combinations in a particular language (Castles et al., 2018, p. 19). Eventually "spellings, pronunciations and meanings of specific words in memory" are bound together and written word meaning is accessed directly in memory rather than indirectly through decoding (Ehri, 1999, pp. 88-96; Ehri 1992, p. 108).

1.2.2. L1 phonological decoding and word recognition

Phonological decoding is also believed to influence how representations of written words are formed/stored in memory, and it remains accessible even in skilled readers (Castles et al., 2018). Share (1999) suggests that phonological decoding is "item-based" or depends on frequency of exposure to, and active decoding of, written forms (p. 96). In other words, phonological decoding supports the learning of written forms of words throughout all stages of reading development (The Self Teaching Hypothesis; Share, 1999). Phonological decoding of embedded non-words whilst reading connected texts was found to support children's word recognition and spelling (Share, 1999; Ricketts, Bishop, Pimperton, & Nation, 2011). The process of actively decoding the written form seems to lead to improved "lexical quality" and the "precise" yet "flexible" mental storage of written forms (Castles et al., 2018, p. 21). Kaefer (2016) too posits that phonological decoding activates explicit awareness of orthographic knowledge (spelling, word recognition) in early readers (mean age 5 years 10 months).

1.3. Learning to read in a first language (L1): Skills that underpin phonological decoding

In an L1, word recognition involves "turning listeners into readers" (Stuart, Masterson, & Dixon, 1999, p. 100), and the phonological decoding processes outlined previously are predicated on learned skills involving "lower level verbal processing mechanisms" (Koda, 1992, pp. 52-57) such as phonological awareness, phonological recoding, and verbal working memory (McBride Chang & Ho, 2005, p. 119). Phonological awareness is "one of the best predictors of literacy outcomes" (Henbest & Apel, 2017, p. 304) and involves the ability to segment and manipulate spoken words (Verhoeven, 2011). This sensitivity involves skills such as syllable identification, the ability to manipulate (delete/exchange) phonemes in spoken words, and sensitivity to onset-rime, or sounds before and after the nucleus of a syllable (Antony, Lonigan, Driscoll, Phillips, Burgess, & Cantor, 2002, pp. 68, 87). These skills are deemed essential in recognising how sounds map onto their respective letters (Ehri, 1999, p. 102). Morphological awareness or the ability to identify, reflect on, and manipulate the smallest units of meaning are also an important factor in word recognition, spelling, and reading comprehension (Apel, Wilson-Fowler, Brimo, & Perrin, 2012; Deacon & Kirby, 2004).

1.4. Teaching word recognition: The role of phonics in L1 English reading instruction

Phonics instruction is currently practised in many English L1 school education contexts (England, USA, Canada, Australia, New Zealand) as a means of developing foundational reading skills through phonological decoding (Department of Education, Science & Training, 2005; National Reading Panel, 2000; The Rose Review, 2006). Whilst phonics-only approaches to teaching early reading remain contentious (Wyse & Styles, 2007), there is evidence that teaching approaches which combine phonics or code instruction (sound/letter links) with meaning-focused activities (understanding text) resulted in faster reading progress in primary school children (Adams, 1990; Chall, 1967). Meta-analyses of experimental research have found that systematic phonics instruction which promotes the learning of sound/letter links in targeted and sequenced programmes, had a statistically significant effect on reading accuracy (Ehri, Nunes, Stahl, & Willows, 2001; Torgerson, Brooks, & Hall, 2006), reading comprehension (Ehri et al., 2001), and spelling (Ehri et al., 2001).

Commercial phonics packages such as Jolly Phonics (Lloyd & Wernham, 1992) or Read Write Inc. Phonics (Miskin, 2011) promote pedagogic routines involving production of isolated phonemes to segment sounds in spoken words, picture mnemonics to memorise letters and exemplar words which model the target sound/spelling links. These activities are supported by research which showed that articulatory gestures supported memorisation of sound/spelling links and word reading performance when coupled with practice in segmenting words (Boyer & Ehri, 2011) and that letter names supported by picture mnemonics lead to faster learning of sound/spelling links (Shmidman & Ehri, 2010). However, as phonics instruction "instils"

rather than "installs" the code (Gough & Wren, 1999, p. 75), it is essential that phonics-related skills are operationalised through reading aloud which involves applying sound/spelling links knowledge (Hatcher, Hulme, & Ellis, 1994). Recall that this is likely to lead to fine-quality lexical representations stored in the learner's memory (Castles et al., 2018). It is further argued that frequent and diverse read aloud texts support word recall and new word spelling (Rosenthal & Ehri, 2011).

To conclude, systematic phonics instruction supports word recognition and is therefore recommended by researchers as a valuable pedagogic tool in L1 literacy development (Castles et al., 2018). It is also important to note that, as the ultimate goal of any reading instruction programme is reading comprehension, secure and automatized word recognition is only one part of a complex developmental process. Other factors likely to influence reading comprehension outcomes include oral language, parsing, inferring, comprehension monitoring, exposure to print, morphological awareness, reading motivation, and cognitive resources such as working memory (Castles et al., 2018; Perfetti & Stafura, 2014).

1.5. Learning to read in an L2
1.5.1. A repeated process – shared, underlying skills and knowledge
Learning to read in a second or additional language is a "repeated process" (Koda, 2008, p. 74) and involves a constantly active "two language processing system" (Grabe & Stoller, 2002, p. 35). Research in bi-/multi-lingual settings has determined that some foundational L1 decoding related skills are evident in L2 literacy development. L1 phonological awareness supported children's L2 reading development (Durgunoğlu, Nagy, & Hancin-Bhatt, 1993) and predicted L2 word reading skills (Lesaux et al., 2008). Opinions vary as to whether these L1 resources are available in the early stages of L2 oral development (Genesee, Geva, Dressler & Kamil, 2008, p. 68) or whether a particular level of L2 competence is required (Interdependence Hypothesis; Cummins, 1979).

However, it is important to note that differences in orthographies and syllabic structures as well as participant age are likely to play a role in cross-linguistic transfer of L1 literacy skills (Genesee et al., 2008). For example, Goswami (2006) posits that readers of L1 English learn to map sounds with letters in chunks (Psycholinguistic Grain Size Theory) due to the consistency of English at rime level. On the other hand, L1 readers of more transparent orthographies such Spanish, with consistent 1:1 mappings between sound and print, will have recourse first to decoding at the individual phoneme level (Landerl, 2006, p. 518).

Interestingly, in FL settings, a bi-directional influence for foundational L1 literacy skills has been noted. Clusters of L1 skills such as spelling, word decoding, phonological awareness, vocabulary, and reading comprehension influenced FL attainment and aptitude in high school (Sparks & Ganschow, 2012; Sparks, Patton, Ganschow, & Humbach, 2009). These researchers had earlier identified that similar L1 skills acted as a foundation for L2 proficiency, arguing that weak phonological processing could negatively influence both L1 and L2 learning (Linguistic Coding Deficit Hypothesis; Sparks & Ganschow, 1991). Research in younger FL classrooms has evidenced bi-directionality of phonological awareness influence, finding that L2 phonological awareness can also support the development of L1 phonological awareness in learners aged 7-8 (Murphy, Macaro, Alba, & Cipolla, 2014).

1.5.2. A different process- new sounds and new mappings
Word recognition and phonological decoding rests upon phonological awareness (deemed transferable) but also on orthographic mapping or the development of links between sounds and their respective letters (Ehri, 2014). In an L1, these sounds are already secure and just need to be mapped to letters. However, in an L2 or FL, sounds and letters will either need to be learned "from scratch" (i.e., new FL sounds linked to new FL letter/s) or re-mapped (i.e. existing letters/sounds linked to different counterparts).

In terms of remapping, Koda (2007) argues for a process where L1 knowledge and skills are transformed into additional L2 competencies through exposure to L2 print. However, this seems problematic bearing in mind that L1 readers' knowledge and decoding skills are likely to have transformed from effortful to automatized through the process of learning L1 reading (Castles et al., 2018; Ehri, 2014; 1999). The task for learners of L2/FL re-mappings will likely involve adapting automatized L1 sound/spelling mappings, and automatized cognitive language learning mechanisms are believed to be particularly resistant to adaptation (Mitchell, Myles, & Marsden, 2019).

Although there is little evidence concerning the learning of novel phonology in FL settings, L2 phonological acquisition has been extensively investigated. Whilst recognising important differences between L2 and FL contexts in terms of input quality, extent, and intensity, theories relating to L2 phonology

acquisition could give some insight into the process of learning novel sounds when an L1 phonological system is well established.

Opinions differ regarding ease of acquisition of L2 phonology. For some, novel sounds are deemed harder to learn either due to entrenched L1 phonological representations or because they require the creation of new L2 sound categories (Escudero & Boersma, 2004; Lado, 1957). Others assert that dissimilar sounds are easier to learn as they are more likely to be perceived and therefore result in "new L2 phonetic categories" (Speech Learning Model; Flege, 1995). Crucially, L2 sounds will differ from monolingual sound acquisition as L2 sounds will interact with the L1 and presumably subsequent, additional languages over an individual's lifespan (Flege, 1995; Nimz, 2018).

For example, French nasal sounds, for example, vowels (often novel for learners) have proved problematic showing variation across advanced learners and particular types of nasal vowels: <ON> /ɔ̃/was the most accurately produced but <IN> /ɛ̃/ was the least accurately produced and tended to be substituted by <AN/EN> /ɑ̃/ (Adeline Charlton, 2014; Neveu et al., 2010). Foreign language instruction seems to assist the development of novel French nasal vowels and shifts in phonological representations occur between proficiency levels (Martinez, 2016). Interestingly, once learners are literate in their L1, orthographic knowledge influences the perception and production of L2 phonology (Bassetti & Atkinson, 2015; Escudero & Wanrooij, 2010; Nimz, 2018). L1 orthography is indiscriminately activated when written cues are involved in sound perception and this has an ambivalent effect on perception (Escudero & Wanrooij, 2010). It should also be noted that other factors, such as L1 syllable structure, L1 phonotactics, and individual differences, influence L2 sound perception, particularly of vowels (Tyler, Best, Faber, & Levitt, 2014).

1.6. Teaching FL/L2 reading

Bearing in mind issues such as transfer of L1 related skills and competencies, the instructional task for teachers differs from that of teaching children to read in an L1. Assumptions should not be made that pedagogic tools and approaches relevant in the L1 will automatically have the same, or indeed any, effect on FL development.

Clearly when learning to read in a second or subsequent language, the learner is not a *tabula rasa* (Porter, 2014). Firstly, a learner approaches L1 reading with a fairly extensive L1 spoken vocabulary, internalised L1 grammar, and L1 phonological inventory. The 6-7 year old (grade 2) native speaker has an average L1 root word vocabulary of 5200 words, rising to 8400 words by age 11 (Biemiller & Slonim, 2001). In instructed L2/FL contexts, the starting point is somewhat different. Often, literacy instruction in the FL/L2 languages classroom is initiated after the development of oral skills and it is suggested that L2 phonological decoding has limited effectiveness unless the oral representation of the word is already known (Grabe & Stoller, 2002, p. 42). However, in FL settings where instructional time is limited, waiting for the right or optimum time to introduce decoding and focused literacy instruction may not be a pedagogical option. On the other hand, it appears as though the essential foundational literacy skills of phonological awareness may be available to learners if they are already developed in the L1. It is also likely that increasing cognitive maturity of children in the later stages of primary school (when FL instruction generally commences) may assist in the teaching and learning of FL literacy.

To summarise, word recognition is a key skill in developing FL literacy and is underpinned by the knowledge and application of FL sound/spelling links. Research indicates that teaching FL word recognition is likely to be a shared and yet different process to its development in the L1. Already-literate learners could bring transferable skills to the task of FL word recognition, such as phonological awareness, relative cognitive maturity, and shared sound/spelling links. However, some L1 skills—sound perception, knowledge of FL sound/spelling links, and FL oral vocabulary—are likely to complicate the learning process. It is therefore important that, in primary FL settings, research is conducted to explore the potential effectiveness of L1 literacy pedagogic practices and curricular expectations.

2. Method

The study set out to investigate whether/how children studying French as a new FL made progress at reading aloud FL word cards during and after a 23-week teaching programme which comprised systematic and explicit phonics instruction in French (research question 1). The test measures also aimed to contrast how successfully the children read aloud known (familiar) words that they had practised every week with how they decoded novel (unfamiliar) words (research question 2). Lastly, the study aimed to evaluate

whether there were any relationships between L1 literacy measures, FL proficiency scores, and reading aloud performances at each test time (research question 3). These analyses would account for influential factors in the learning process and therefore possibly explain some of the variation in FL reading aloud outcomes.

- RQ1: To what extent do learners engaged in systematic and explicit FL phonics instruction make progress in reading aloud familiar and unfamiliar FL words at mid-, post-, and delayed post-test?
- RQ2: Are performances on reading aloud familiar and unfamiliar word cards different at mid-, post-, and delayed post-test?
- RQ3: To what extent are L1 literacy measures and FL global proficiency scores related to FL reading aloud proficiency at each test time?

2.1. Sample

This study took place in two English primary schools. The total sample (n = 45) was tested at three main test points: pre-, post-, and delayed post-test on reading aloud in sentences. The focus of this paper concerns a sub-sample of students (n = 23) who were tested on all the main measures plus word card reading aloud at mid-test (approximately 10 teaching weeks after the phonics instruction commenced), post-test (23 weeks after), and delayed post-test (7 weeks after the teaching programme finished). The mean age of the sub-sample was 10.07 years (minimum 9.10, maximum 11.0) and the children were taught in mixed-year classes (England Years 5 & 6; US Grades 4 & 5). All of the children spoke English as their L1. Although some had exposure to other languages at home (e.g. Nepalese, Arabic, Dutch), these were not spoken as primary home languages. Both schools were not in areas of social deprivation and were situated in small, rural villages. The FL provision in both settings was identical: the same French teacher (the researcher) taught one 50-minute FL lesson per week. This teacher/researcher had worked in School 1 for 2 years and School 2 for 6 years prior to the research.

It should be noted that the mixed-year class in School 1, in particular those children in Year 6, had had one year less cumulative instruction than the Year 6 children in School 2. No children had previously received systematic and explicit French phonics instruction, and FL literacy generally had not been given special or sustained emphasis. Nonetheless, there was only one significant difference in FL pre-test scores between schools, which concerned FL reading aloud scores. Parents and guardians were approached for informed consent, and the children's verbal consent was sought in order to conform to ethical requirements for school-based teaching and learning research. This was an action research study so there was flexibility within the design to adapt and amend teaching practices. However, in order to support robustness and generalizability/transferability, the principles underpinning the overarching teaching programme (of which FL phonics instructed formed a component) and those principles underpinning the specific strand of phonics instruction are outlined below.

2.2. Pedagogic principles for literacy instruction

This small-scale study aimed to build a phonics instruction programme which could assist primary school learners of French as a FL, with extremely limited input, to start to develop FL sound/spelling links. The teaching programme included a weekly phonics component which formed part of lessons which aimed to develop both FL oracy and literacy skills. The phonics instruction specifically had three overarching principles based on the extant literature regarding L1 phonics instruction.

2.2.1. Principle 1: Focus on FL sounds

First, it was crucial to develop distinctive FL sounds (phonology) as phonics instruction involves linking known sounds to letters. This task required the development of distinctive new or adapted phonemes such as nasal vowels. Rather than aiming for native-like pronunciation, distinctive FL sounds were evidenced by sounds which were clearly differentiated in production from their closest L1 counterparts.

There are particular French phonemes which do not have exact counterparts in the L1 (e.g. nasal vowels /ɔ̄/, /ā/, /ɛ̄/, /œ̄/). Furthermore, some French phonemes have similar English counterparts and could therefore be difficult to discriminate (e.g. French front rounded vowel /y/ versus the similar English back rounded vowel English /u:/). Bearing in mind the L2 phonological development literature, the pedagogic programme therefore comprised a small element which focused on FL sound perception (discrimination) and production.

2.2.2. Principle 2: Systematic and explicit phonics instruction

In line with commercial phonics packages (e.g., Read Write Inc. or Jolly Phonics) sounds and letters were taught explicitly and in a planned order (Table 1). Less attention was given to phonological awareness such as the ability to segment spoken words into sub-lexical units—individual phonemes and/or onset-rime or syllables—as this knowledge was assumed to transfer from the L1 (e.g., Genesee et al., 2008; Lesaux et al., 2008). However, children were occasionally asked to segment spoken words into syllables and to segment and blend written words.

Table 1
Target FL sound/spelling links

	Sound	Letters	Exemplar	Gesture	Visual	Links with L1
Pre-lesson	Letter names	Alphabet				
Week 1	Final silent letters	PETS, <P>, <E>, <T>, <S>				
Week 2-5	/ɔ̃/	<ON>	Cochon *(pig)*	Finger on nose - snout	Picture of pig	Novel FL sound/existing L1 grapheme
	/y/	<U> and <Û>	Putois *(ferret)*	Fingers pinching nose (smell)	Picture of ferret	Novel FL sound/ existing L1 and novel FL grapheme
	/j/	<LL>	Billet *(ticket)*	Taking a ticket	Picture of train ticket	Existing L1 sound/existing but different L1 grapheme
Week 6	/ɑ̃/	<AN> and <EN>	Enfant *(child)*	Rocking a baby	Picture of baby	Novel FL sounds/existing L1 grapheme
	/ɛ̃/	<IN>, <IEN>, <EIN>	Singe *(monkey)*	Arms bent at sides	Picture of monkey	Novel FL sounds/existing L1 and novel FL graphemes
Weeks 7-11	Practice all previous sound/spelling links					
Week 12	/e/	<E>, <ER>, <EZ>, <ET>	Éléphant *(elephant)*	Arm imitating trunk	Picture of element	Novel FL sound/existing L1 and novel FL graphemes
	/o/	<EAU>, <O>, <Ô>, <AU>	Océan *(ocean)*	Arm moving like a wave	Picture of ocean	Novel FL sound/existing L1 and novel FL graphemes
Week 13-20	Practice all previous sound/spelling links					

Sounds and their corresponding letters were also presented through an exemplar word, e.g. cochon, which aimed to allow for associative learning and the possibility that this could be used through analogy to decode the same grapheme in unfamiliar words. It was anticipated that the learning of FL sound/spelling links would be facilitated by shared links with English but that learning to adept/amend existing mappings or

creating entirely novel ones would be more problematic. The sounds and their links were taught on a cumulative basis so that, as a new sound was introduced, previous sound/spelling links were still practised regularly.

2.2.3. Principle 3: Applying sound/spelling links through reading aloud practice

It was also deemed necessary to offer children opportunities to practise applying these sound/spelling links through reading and writing opportunities. Weekly spelling activities involving core vocabulary (not phonics exemplar words) supported the children's learning of links between sounds and letter in French. The children were asked during a sequence of tasks designed to support vocabulary memorisation, to recall a word orally from a visual cue and then to attempt to write this word from memory (having seen the written form several times previously). Wherever possible, core vocabulary was utilised to teach phonics sound/spelling links such as ondule *(slither)* <ON> = /ɔ̃/and fourrure *(fur)* <U> = /u/. It should be noted that these words were chosen to support broader, communicative language use rather than act solely as practice for the phonics element of the instructional programme. In reading aloud opportunities such as stories and pairwork, peer-assessed reading of sentences were used to practise phonological decoding. Again, these stories were mostly focused on meaning-related activities in line with integrated approaches to reading but comprised occasional decoding opportunities (Adams, 1990).

2.3. Procedures

The reading aloud test battery targeted specific, taught FL sound/spelling links as well as a broader range of linguistic skills and knowledge. Due to the lack of existing resources for testing phonological decoding in primary school learners of French, test items were designed by the teacher/researcher. As there was no pilot sample with similar language learning experience in terms of age, language, and extent of input, these tests were not piloted. The broader linguistic FL assessment tasks were closely aligned with classroom activities and designed to test the constructs of FL general proficiency and FL global literacy. These were administered at pre-, post-, and delayed post-test. The post-test was conducted immediately after the instructional programme ended, and the delayed post-test was completed seven weeks later. The general proficiency construct comprised an oral vocabulary recognition test and an elicited imitation test designed to measure developing interlanguage. The FL global literacy tests involved reading aloud target words in sentences (with illustrative pictures to support meaning) and reading for global meaning and detail of a short text. An additional reading aloud test was implemented during the teaching programme in order to capture any development in a) applying sound/spelling links through decoding familiar exemplar phonics words and b) applying these taught sound/spelling links in novel contexts (unfamiliar words). This test was delivered with a sub-sample of pupils (n = 23) at mid-test, post-test and delayed post-test. The whole sample (n = 45) took the FL reading aloud word card test at post-test and delayed post-test.

2.4. Research instruments

2.4.1. FL general proficiency: Receptive vocabulary and elicited imitation

This variable was addressed through two complementary, general proficiency (non-literacy) measures. The *FL receptive vocabulary* test was based on the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) and the ELIAS British Picture Vocabulary Scale (Dunn, Dunn, Whetton, & Burley, 1997). Lexical items which had been taught over previous years where chosen and individual worksheets were completed. From a spoken cue (provided by the teacher), the children had to circle the picture which correctly represented each lexical items. Some items were discrete words (e.g., *le père – father*) and some were formulaic utterances/constructions (*j'adore faire du cheval – I love horseriding*). The pictorial cues were familiar to the children as they were largely derived from historic teaching materials used by the teacher.

The *FL elicited imitation* test was designed to show developing implicit (subconscious) L2 knowledge (Erlam, 2006). Learners listened to utterances incorporating particular syntactic or morphological demands, such as interrogatives or negation, and were asked to repeat them after a short gap. Prior to listening, they were guided to process the utterances for meaning by being warned that they would have to answer questions about the utterances after they had heard them. All the stimuli were designed to be long enough to exceed working memory capacity and therefore could only be reformulated by accessing meaning and re-constructing the utterance from interlanguage. For example, the sentence *Xavier, il n'aime pas aller à l'école* (*Xavier does not like going to school*) was elicited to explore learning of negation expressed through dislike. Elicited imitation is deemed an efficient way of accessing spontaneous oral ability in young learner

classrooms where there are often practical constraints such as the amount of available test time (Campfield, 2015). To measure performance, reproductions were marked by target-like syllables. Each utterance, regardless of length, could achieve a maximum score of 4 and syllabic errors were deducted from this as required (Lonsdale, Graham, Kennington, Johnson, & McGhee, 2008). Therefore, a sentence with two syllable errors (non-target-like reproduction) would receive a mark of 2. Often elicited imitation tests with younger learners are marked on a rating scale such as: 3 = mainly target-like reproduction to 0=unintelligible production. It was felt that a syllabic scoring mechanism might give a better indication of the extent of FL oral processing (reproduction in terms of sentence length and complexity) rather than an arguably more subjective rating scale which focused on evaluation of overall accuracy of sentences. A small sample of post-test scores were second-marked by the same rater with a reliability score of 67%.

2.4.2. FL literacy proficiency: Reading aloud (sentences) and reading comprehension

For the *FL reading aloud (text)* test, short sentences illustrated with pictures were read aloud by each child. For example, *Monsieur Laurent travaille au marché* (*Mr Laurent works on the market)* was constructed to elicit reproduction of specific, underlined sound/spelling links. Individual words (containing instructed FL sound/spelling) were scored for target-like recoding, and a clear scoring scheme was set out to contribute to rater reliability. Recoding was strictly marked to reflect explicitly taught characteristics of FL sounds and their links to letters. In other words, recoding of <ON> was only deemed target-like (scoring 1 = correct) if it reflected nasality /ɔ̃/. Even though /ɒn/ would have been largely intelligible to a sympathetic speaker of the language, it was marked incorrect as production/discrimination of novel FL phonology was an integral part of the instructional programme. Second marking by the same rater yielded a reliability score of 83%. This test was conducted on a one-on-one basis, during lesson time in a private setting at each school.

The *FL reading comprehension* test aligned with primary FL learning outcomes including: "understand the main points and some detail from a short written passage" and "reading texts for enjoyment or information" (DfES, 2005, p. 57). A short passage with 11 (8%) unfamiliar words was designed as a police report concerning a robbery. Questions were presented in children's L1 and designed to test reading for understanding of the main points and some detail. Responses were designed to require "limited production" (McKay, 2006, pp. 106, 186) in that children could identify the meaning by selecting a picture. For instance, for the question "What item was stolen?" children could choose from pictures of a wallet, a bag and some house keys. It was anticipated that this would mitigate cognitive loading effects for children with weaker L1 literacy. Scoring was dichotomous, and this test was conducted using individual worksheets in a whole class setting without conferring. Children were allowed approximately 15 minutes to complete the test.

### 2.4.3. FL phonics: Read aloud word card

This test was designed in line with action research principles and cycles of research where initial investigations yield new or more refined lines of inquiry. The teacher/researcher deemed, during the instructional programme, that learning and applying exemplar sound/spelling links was challenging for the children. She felt it would be interesting to uncover how progress was unfolding with respect to decoding exemplar phonics words and also whether this knowledge was being applied to novel contexts. With this in mind, illustrated word cards were designed (arguably more as a formative assessment tool than as a research instrument). At mid-test, due to time constraints, a small number of pupils (n = 23) were tested. The children were selected based on their FL reading aloud performance at pre-test (sentence) from each third of the scoring range (Table 2). The children were each asked to shuffle the cards and then read aloud each word whilst being audio-recorded. Each child sat alone in the corridor outside the classroom at mid-test whilst the teacher taught the rest of the class. No feedback was given at any stage. At post-test and delayed post-test, this word card measure was incorporated into the overall test battery and delivered on a one-to-one basis with the teacher/researcher in a private room. Each word card had a picture for meaning (Figure 1). For every targeted FL sound/spelling link there was a card with the phonics exemplar word and card with the same grapheme, in a different position, in an unfamiliar word.

Table 2
L2 reading aloud pre-test, quartile breakdown

Pre-test read aloud scoring (max 10)	Participants
Low (0-2)	8
Mid (3-4)	10
High (5-6)	5
Total participants	23
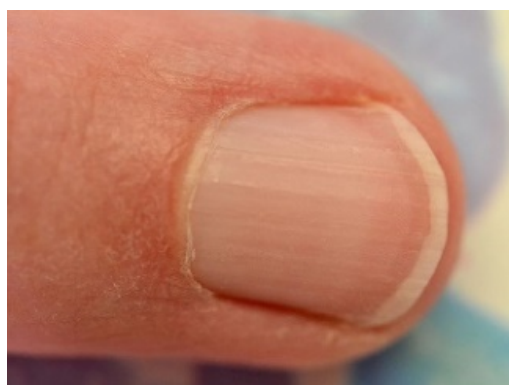
Cochon   Ongle

Figure 1. Example of exemplar phonics word card

2.4.4. Underlying L1 literacy measures

In line with national data collection in English primary schools at the time, this study had access to a range of standardised L1 literacy test scores (reported on a national level by school and child) and classroom-based informal teacher assessments. These were deemed potentially important in explaining the influence of L1 literacy on FL reading aloud performances. Children's chronological reading ages were collected using a standardized New Group Reading Test (GL Assessment, 2018). This test comprises two parts: sentence completion (decoding with some comprehension) plus passage comprehension where children are tested on a range of comprehension skills. The Single Word Spelling Test (GL Assessment, 2018) was also administered by the regular class teachers to determine children's chronological spelling age through spelling words in context (sentences). This is also a standardized, commercially produced test.

In addition to these standardized tests, the children's regular class teachers (not the researcher who was their longstanding yet visiting FL teacher) undertook ongoing formative assessments of each child's progression in all curricular subjects. These teacher-centred assessments, backed up by observation and other classroom data such as examples of children's work, were then linked to levels of attainment which at the time formed part of the National Curriculum. Whilst levels have now been eliminated formally from government documentation, they are still sometimes referred to by teachers as a means of benchmarking pupils' performances against their peers and previous cohorts. National Curriculum levels ranged from 1-6 with the idea that, by the end of primary school, expected attainment for children would reach level 4. Each level was accompanied by descriptors for each subject and had sub categories (1c was the bottom sub-level within level 1 whilst 1a was the top sub-level). To facilitate statistical analysis, these levels were converted to a continuous scale e.g. level 1c was 1 and level 6a was 18. Numerical levels were assigned to each child for L1 reading (reading aloud, comprehension) and L1 writing (spelling, complexity, genre, cohesion).

3. Results

A Shapiro-Wilk test (as *n* equalled <50) showed that assumptions of normality were violated for 33% of the pre- and post-test FL constructs. Given that result, and bearing in mind the relatively small sample size, subsequent statistical analyses were conducted using non-parametric techniques. Mann Whitney U-Tests were conducted to identify whether there were any differences between sub samples in each school for L1 and FL read aloud measures at pre-, post-, and delayed post-test times. No differences were found between school or gender.

3.1. To what extent do learners engaged in systematic and explicit FL phonics instruction make progress in reading aloud familiar and unfamiliar FL words at mid-, post-, and delayed post-test?

Sections 3.1.1 and 3.1.2 will report on whole-sample performances for reading aloud familiar and unfamiliar word cards, during and immediately after the instructional programme (at mid-test and post-test). Descriptive statistics will portray the extent and nature of developing of FL reading aloud scores at both test times. Sections 3.1.3 and 3.1.4 will then use statistical techniques to establish whether any improvements were statistically significant in reading a) familiar word cards and b) unfamiliar word cards.

3.1.1. Reading aloud familiar and unfamiliar word cards at mid-test

By mid-test, which was conducted 10 weeks after the FL phonics instruction commenced, descriptive statistics showed that progress reading familiar word cards was slow overall (mean score 1.13; median score 1.00) out of a maximum score of 5 (Table 3).

Table 3
Descriptive statistics, mid-test, reading aloud word cards (N = 23)

	Mid-Test							
					Range			
	Mean	**Median**	**Mode**	**SD**	**Min**	**Max**	**Skewness**	**Kurtosis**
Familiar words (max 5)	1.13	1.00	.00[a]	1.100	0	4	1.096	1.168
Unfamiliar words (max 5)	.3042	.00	.00	.653	0	2	1.994	2.934

Note. [a] = multiple modes

Positive skewness and kurtosis values indicated that scores at this test time were clustered towards the lower end of the scoring range (Figure 2). However, the range of scores showed that whilst one child (out of the 23 participants) had scored close to full marks, others had made no progress (seven children scored 0) reading familiar words. Despite weekly lessons which involved practising production and discrimination of FL sounds, FL phonological decoding (spellings to sound), and reading aloud of familiar (exemplar) words, learners had made little progress at reading aloud the practised words mid-way through the teaching programme.
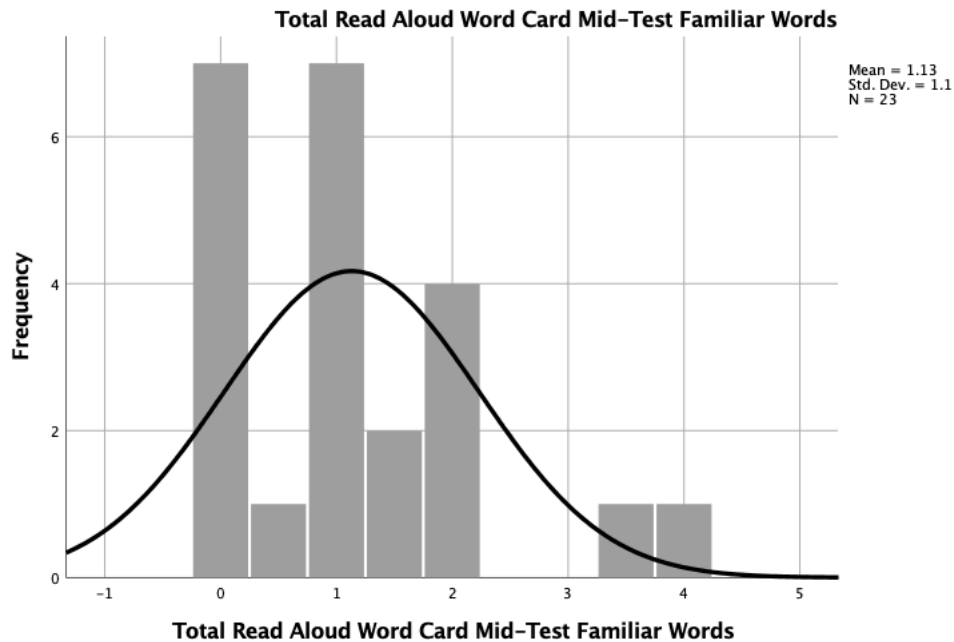
Total Read Aloud Word Card Mid−Test Familiar Words

Mean = 1.13
Std. Dev. = 1.1
N = 23

Figure 2. Distributions of read aloud familiar words at mid-test

Unfamiliar word card reading scores were less variable (range 0 to 2). Eighteen children scored zero whilst two children scored 2 marks at mid-test points. Almost no progress was evidenced at mid-test for reading unfamiliar word cards (unfamiliar mean .3043; median .00) and again skewness and kurtosis showed scores weighted toward the lower end of the range (Figure 3). Perhaps unsurprisingly, bearing in mind the limited progress for familiar (exemplar) words, the children had also made scant progress at applying learned sound/spelling links to novel contexts.
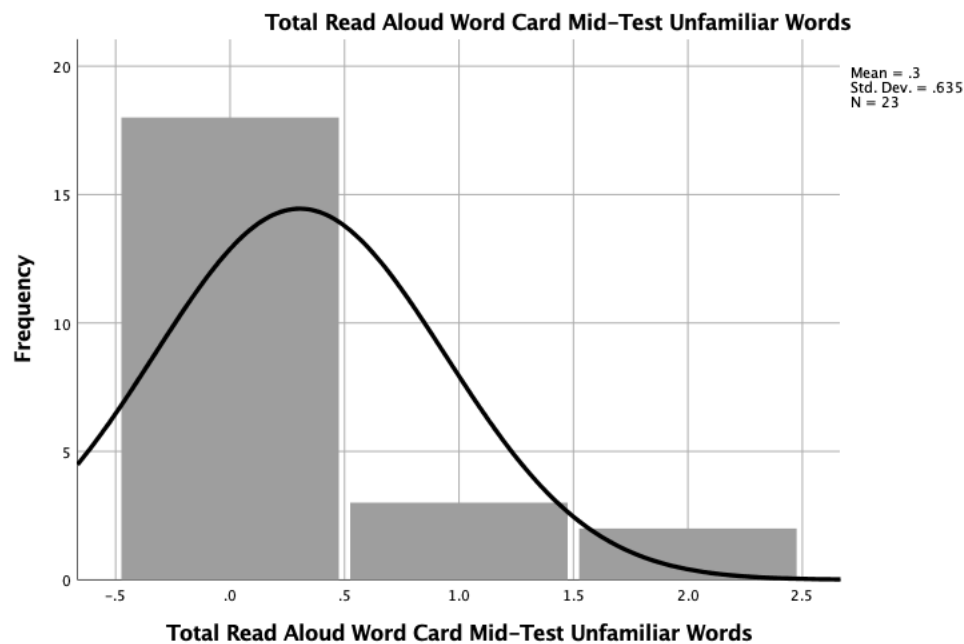
Total Read Aloud Word Card Mid−Test Unfamiliar Words

Mean = .3
Std. Dev. = .635
N = 23

Figure 3. Distributions of read aloud unfamiliar words at mid-test

3.1.2. Reading aloud familiar and unfamiliar word cards at post-test

By post-test, at the end of FL phonics instruction (23 weeks in total), descriptive statistics showed that progress reading familiar word cards had improved (mean score 2.174; median score 2.00) out of a maximum score of 5 (Table 4).

Table 4
Descriptive statistics post-test reading aloud word cards (N = 23)

| | Mid-Test | | | | | | | |
| | Mean | Median | Mode | SD | Range | | Skewness | Kurtosis |
					Min	Max		
Familiar words (max 5)	2.174	2.00	1.00	1.230	0	4.5	-.040	-.708
Unfamiliar words (max 5)	.9130	1.00	.00	.950	0	3	.535	-.934

A change in skewness values (now negative) indicated that scores were clustered towards the mid or higher end of the score range. However, positive kurtosis values indicated that there remained a heavy tail of scores in the lower marks (Figure 4). Two children scored 4/4.5 and two children scored 0 reading familiar words. In other words, the range was still wide between top and bottom scores but the lowest scorers were less numerous. After 23 weeks, the sample seemed to be making progress in reading aloud familiar (exemplar) words.

However, scant progress was evidenced at post-test for reading unfamiliar word cards (mean .9130; median 1.00) and again skewness and kurtosis showed scores tended towards the lower end of the range (Figure 5). Unfamiliar word card reading scores showed slightly different variability (range 0 to 3 maximum) with 10 children scoring zero and 1 child scoring the top mark of 3 at post-test. 23 weeks after the weekly phonics instruction commenced, most children were still unable to consistently apply their developing knowledge of FL sound/spelling links to novel contexts
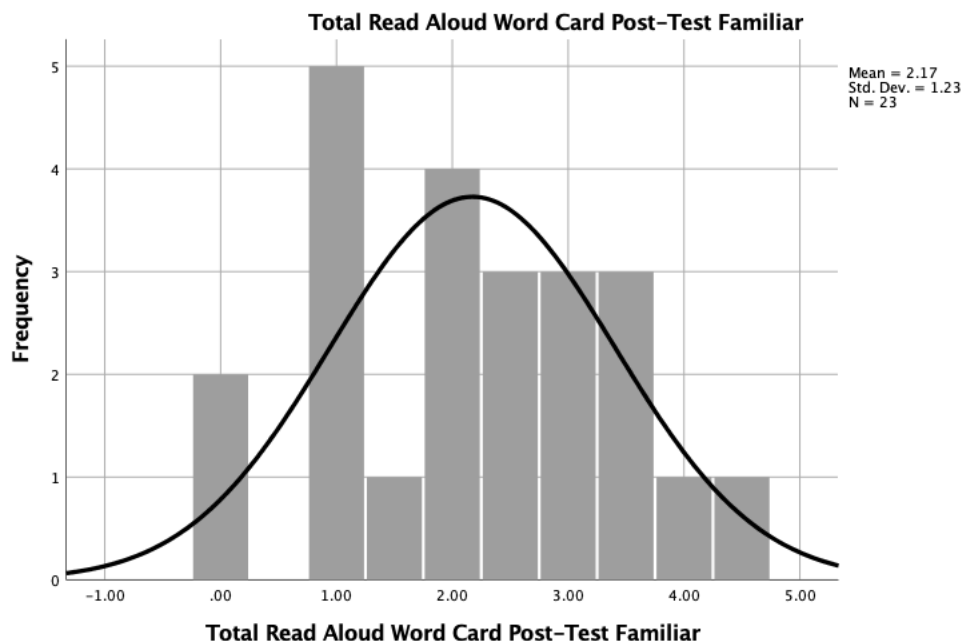

Total Read Aloud Word Card Post–Test Familiar

Mean = 2.17
Std. Dev. = 1.23
N = 23

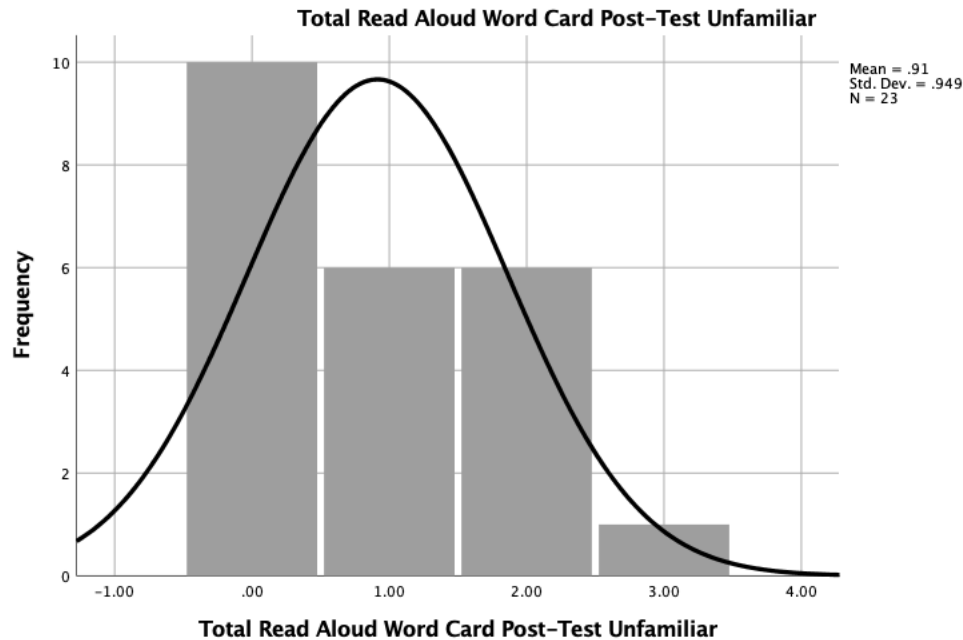*Figure 4.* Distributions of read aloud familiar word cards at post-test

Figure 5. Distributions of read aloud unfamiliar word cards at post-test

3.1.3. Reading aloud familiar word cards: Differences between mid-test and post-test

In this and the following section, scores between mid- and post-test are compared for both kinds of word cards separately to evaluate whether reported improvements were of statistical significance. Wilcoxon Signed Rank Tests (Table 5) show that the scores for familiar reading aloud word cards between mid- and post-test were significantly different. Target-like decoding at post-test was significantly higher (median 2.000) than at mid-test (median 1.000), z = 3.533, p = <0.0001, r = 0.52, moderate effect size. So by the end of the teaching programme, there was a significant improvement in children's familiar word reading aloud scores.

Table 5
Mid-test To Post-Test Familiar FL reading aloud word cards – Wilcoxon Signed Rank Tests (N = 23)

| | Median | | | | | |
	Mid-test	Post-test	z	p	S.E.	Effect Size (Cohen's *d*)
FL Word Card Familiar Reading Aloud	1.000	2.000	3.533	<.0001	19.245	0.52

Note. Max score familiar words 5; unfamiliar words 5.

3.1.4. Reading aloud unfamiliar word cards: Differences between mid-test and post-test

Wilcoxon Signed Rank Tests (Table 6) show that unfamiliar reading aloud word cards between mid- and post-test were significantly different. Target-like decoding at post-test was significantly higher (median 1.000) than at mid-test (median .000), z = 2.617, p = .007, r = 0.40, moderate effect size. This meant that the children became significantly better at reading aloud unfamiliar word cards between the mid and end point of the teaching programme. It is, however, important to note that medians and means for both kinds of word cards at mid- and post-test were still below the mid-point of the scoring range.

Table 6
Mid-test To Post-Test Unfamiliar FL reading aloud word cards – Wilcoxon Signed Rank Tests (N = 23)

| | Median | | | | | |
	Mid-test	Post-test	z	p	S.E.	Effect Size (Cohen's *d*)
FL Word Card Reading Aloud	.000	1.000	2.617	.007	10.937	0.40

Note. Max score familiar words 5; unfamiliar words 5.

Sections 3.1.5 will now turn to performance at delayed post-test approximately 7 weeks after all FL instruction had finished. This will evaluate the durability of any learning and explore whether there was any observed attrition related to word reading aloud in both familiar and unfamiliar contexts. Descriptive statistics will first explore the nature of change (Section 3.1.5), and inferential statistics will then compare performances between post- and delayed post-test (familiar and unfamiliar words treated separately) to evaluate whether any observed differences were statistically significant (Sections 3.1.6. and 3.1.7.).

3.1.5. Reading aloud familiar and unfamiliar word cards at delayed post-test

By delayed post-test, which was conducted 7 weeks after the FL phonics instruction ended (with no French language input in between), descriptive statistics show a dip in means (to 1.8696 out of a maximum of 5) but not in medians (Table 7).

Table 7
Descriptive statistics delayed post-test reading aloud word cards (N = 23)

| | Mid-Test | | | | | | | |
| | | | | | Range | | | |
	Mean	Median	Mode	SD	Min	Max	Skewness	Kurtosis
Familiar words (max 5)	1.8696	2.00	.00	1.440	0	4.5	.136	-1.093
Unfamiliar words (max 5)	.6957	.00	.00	.974	0	3	1.331	.876

A reversal in skewness scores (Table 7, Figure 6) was also noted which illustrated that scores had shifted from the higher range (at post-test) back to the lower range (by delayed post-test). This confirmed that reading aloud of familiar FL word cards had suffered attrition once the weekly FL instruction (phonics and language) had ceased.
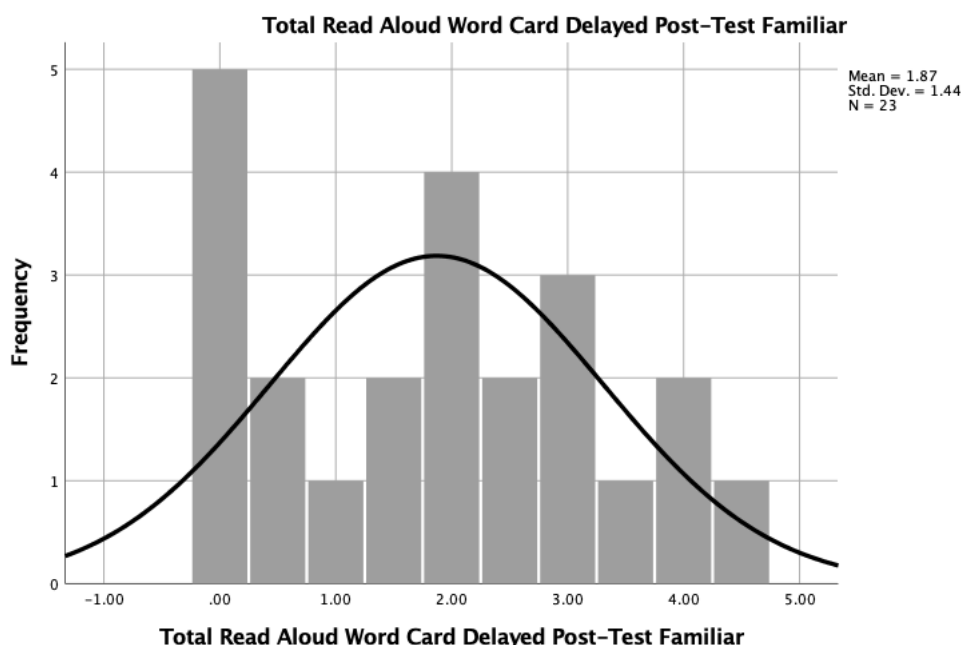

Figure 6. Distributions of familiar read aloud word cards at delayed post-test

For unfamiliar word cards the mean had decreased slightly, from an already low point at post-test. Interestingly, the range of scores for familiar and unfamiliar words were identical to those at post-test and the skewness and kurtosis values remained positive (Figure 7). It seemed as though progress in reading aloud of sound/spelling links in unfamiliar contexts was resistant to instruction during the programme and therefore scores were not particularly impacted once the FL instruction finished.
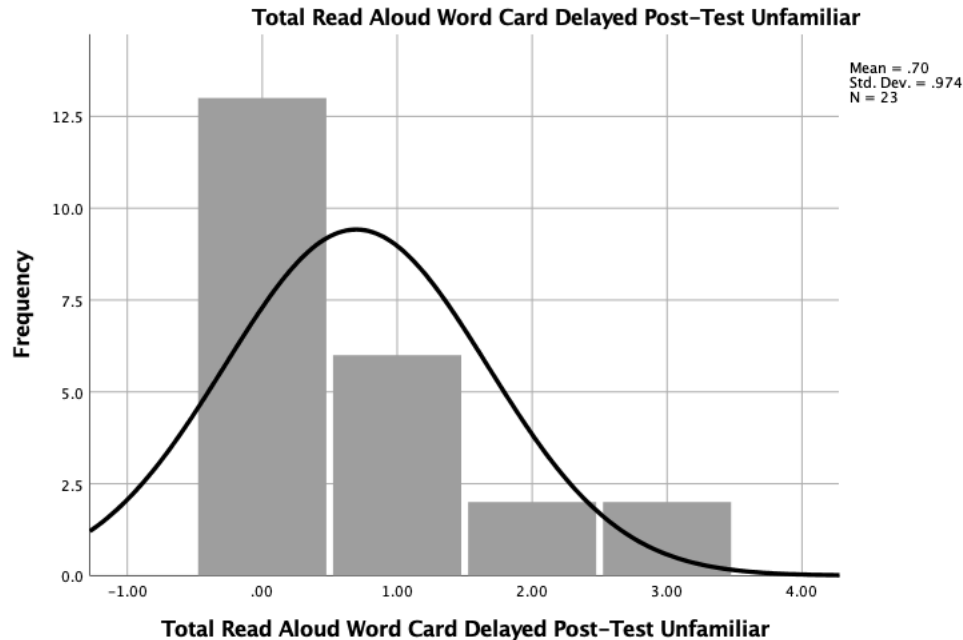
Total Read Aloud Word Card Delayed Post–Test Unfamiliar

Mean = .70
Std. Dev. = .974
N = 23

Total Read Aloud Word Card Delayed Post–Test Unfamiliar

Figure 7. Distributions of unfamiliar read aloud word cards at delayed post-test

3.1.6. Reading aloud familiar word cards: differences between post-test and delayed post-test

Wilcoxon Signed Rank Tests were conducted to evaluate the extent of attrition in students' scores on word card read aloud tests (Table 8) with a Bonferroni correction applied for repeated tests (p = .0125). For familiar word card reading between post-test and delayed post-test (Table 8), target-like decoding clearly was not significantly lower at delayed post-test (median = 2.000) than familiar word card reading at post-test (median = 2.000), z = -1.613, p = .107. In other words, there was no significant attrition between post-test and delayed post-test for children reading aloud familiar words.

Table 8

Post-Test to Delayed Post-Test Familiar FL reading aloud word cards – Wilcoxon Signed Rank Tests (N = 23)

	Median					
	Post-test	**Delayed post-test**	**z**	**p**	**S.E.**	**Effect Size (Cohen's d)**
FL Word Card Reading Aloud	2.000	2.000	-1.613	.107	20.767	n/a

Note. Max score familiar words 5; unfamiliar words 5.

3.1.7. Reading aloud unfamiliar word cards: differences between post-test and delayed post-test

For unfamiliar word card reading (Table 9) between post-test and delayed post-test, target-like decoding was not significantly lower at delayed post-test (median .000) than unfamiliar word card reading at post-test (median 1.000), z = -1.155, p = .248. So, there was also no significant attrition between post-test and delayed post-test for unfamiliar word card reading.

Table 9

Post-Test to Delayed Post-Test Unfamiliar FL reading aloud word cards – Wilcoxon Signed Rank Tests (N = 23)

	Median					
	Post-test	**Delayed Post-test**	**z**	**p**	**S.E.**	**Effect Size (Cohen's d)**
FL Word Card Reading Aloud	1.000	.000	-1.155	.248	12.124	n/a

Note. Max score familiar words 5; unfamiliar words 5.

Overall, then, moderate statistically-significant progress was made for familiar words between mid- and post-test and for unfamiliar words between mid- and post-test. Neither familiar nor unfamiliar words,

examined separately, showed any significant difference between post- and delayed post-test. So the learning of sound/spelling links observed in both familiar and unfamiliar contexts was relatively long-lasting. Nonetheless, it is important to contextualise the nature of the significant differences across test times and test conditions. The shape of the distribution of scores for familiar word cards changed between mid- and post-test but the spread of scores were still centred around the lower scoring range.

3.2. Are performances on reading aloud familiar and unfamiliar word cards different at mid-, post-, and delayed post-test?

Now the analyses will focus on whether word familiarity affects FL reading aloud performances as the descriptive data examined so far suggests this is likely. Sections 3.2.1 to 3.2.3 will take each test-time in isolation and explore whether FL reading aloud outcomes for familiar versus unfamiliar word cards were statistically different. Alongside comparison of group medians, a more detailed examination of performances will be covered in section 3.2.4 at the level of word card.

3.2.1. Reading aloud familiar versus unfamiliar word cards at mid-test

Children's scores were compared, using Wilcoxon Signed Rank Tests, between tasks (familiar and unfamiliar word cards) at each time point. This was to discern whether there was a difference at each test time between reading familiar, exemplar words (rehearsed, whole word pronunciation) compared to unfamiliar words (application of FL decoding and therefore FL sound/spelling link knowledge). Table 10 shows that mid-test familiar word card decoding was significantly higher (median = 1.000) than for unfamiliar word card reading (median = .000), z = -3.368, p = 0.001, r = -0.49, medium effect size. In other words, at mid-test there was a marked difference in overall performance when reading aloud familiar compared to unfamiliar words.

Table 10
Mid-Test Familiar versus Unfamiliar FL reading aloud word cards – Wilcoxon Signed Rank Tests (N = 23)

	Median					
	Familiar	**Unfamiliar**	**z**	**p**	**S.E.**	**Effect Size (Cohen's *d*)**
FL Word Card Reading Aloud	1.000	.000	-3.368	.001	17.367	-0.49

Note. Max score familiar words 5; unfamiliar words 5.

3.2.2. Reading aloud familiar versus unfamiliar word cards at post-test

Similarly, Table 11 shows that at post-test, familiar word card decoding was significantly higher (median 2.000) than unfamiliar word card decoding (median 1.000), z = -3.851, p = <0.001, r = -0.57, medium effect size. By post-test the pattern continued that familiar word learning (reading aloud) was better than unfamiliar word reading.

Table 11
Post-Test Familiar versus Unfamiliar FL reading aloud word cards – Wilcoxon Signed Rank Tests (N = 23)

	Median					
	Familiar	**Unfamiliar**	**z**	**p**	**S.E.**	**Effect Size (Cohen's *d*)**
FL Word Card Reading Aloud	2.000	1.000	-3.851	<.0001	24.668	-0.57

Note. Max score familiar words 5; unfamiliar words 5.

3.2.3. Reading aloud familiar versus unfamiliar word cards at delayed post-test

Finally, Table 12 shows that at delayed post-test, familiar word card decoding was still significantly higher (median 2.000) compared to unfamiliar word card decoding (median .000), z = -3.521, p = <0.001, r = -0.52 medium effect size. This demonstrates that at all test points familiar, exemplar words were more successfully decoded than unfamiliar words.

Table 12
Delayed Post-Test to Familiar versus Unfamiliar FL reading aloud word cards – Wilcoxon Signed Rank Tests (N = 23)

| | Median | | z | p | S.E. | Effect Size (Cohen's *d*) |
	Familiar	Unfamiliar				
FL Word Card Reading Aloud	2.000	0.000	-3.521	<.0001	22.864	-0.52

Note. Max score familiar words 5; unfamiliar words 5.

3.2.4. Reading aloud by word card at each test time

Performances across familiar and unfamiliar word cards at all test times were also examined through analysis of each word card. A breakdown shows that overall performances were centred around specific sound/letter links (Figure 8). Broadly speaking, novel FL sounds linked to existing L1 spellings were learned best in familiar words (*singe*, *cochon*, *enfant*). Those sounds which were either a) similar to an existing L1 sound /y/ or b) an extant L1 sound /j/ which required a different mapping were learned slowest. This was not particularly affected by order of presentation: /j/ as in *billet* was recoded less successfully than /ɛ̄/ as in *singe*, yet /j/ had been presented earlier in the teaching programme. Sounds that were most successfully decoded in post-test familiar conditions were likely to show relative success at post-test unfamiliar conditions, although *singe/câlin* proved an exception. It is important to note that *grand* (unfamiliar) didn't perform as differently as expected although this is likely because the children chose to use this word as part of their extended reading/writing work (the adjective formed part of the core taught vocabulary) and it is also a "false friend" (false cognate) in English.
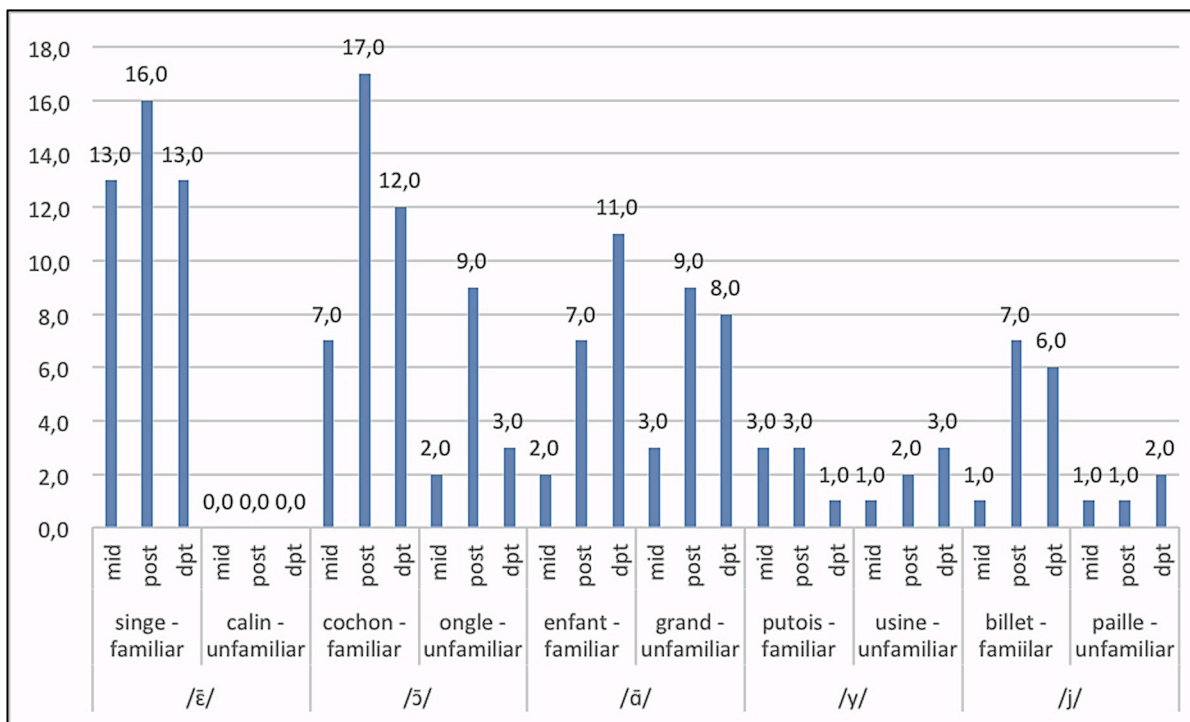

Figure 8. Performance at each test time for specific word cards/sound-spelling links (max score per word/test time 23)

3.3. To what extent do L1 literacy measures and FL global proficiency scores relate to FL reading aloud proficiency at each test time?

This research question seeks to identify whether particular underlying constructs, known to influence FL reading, could explain any of the variance in this study's FL reading outcomes. Firstly, Section 3.3.1 describes the nature of L1 proficiency across the sample and then explores the potential for relationships between L1 literacy measures and FL reading aloud performances at each test time.

3.3.1. L1 literacy measures and FL reading aloud performance

Descriptive statistics (Table 13) show the range of L1 literacy achievement across the sub-sample. Mann Whitney U tests were conducted and confirmed that there were no significant differences in L1 baseline literacy scores between School 1 and School 2.

Table 13

Descriptive statistics for all L1 literacy-related measures

| | Baseline L1 Scores | | | | | | | |
| | | | | | Range | | | |
	Mean	Median	Mode	SD	Min	Max	Skewness	Kurtosis
NGRT L1 reading age	12.709	12.07	13.02	1.675	9.0	15.0	-.003	-.221
SWST L1 spelling age	10.517	10.07	10.02[a]	3.211	7.08	17.0	-1.118	5.115
NC Reading Levels (max 14)	9.96	10.07	10.00	2.383	6	14	2.383	-.319
NC Writing Levels (max 14)	9.35	10.00	10.00	1.991	5	13	-.189	-.325

Note. [a] = multiple modes exist

These data were examined using Spearman's Rho correlation coefficients to establish whether there were relationships between the individual measures of L1 literacy proficiency. Strong positive relationships were found between both National Curriculum reading levels (teacher-assessed) and NGRT standardised reading scores r^s = .809** sig < .0001 (* = .05 alpha; ** = .001 alpha). Moderate relationships were also found between National Curriculum writing levels (teacher-assessed) and SWST standardised spelling scores r^s = .556** sig = .006. NGRT and SWST scores also correlated moderately r^s = .580** sig = .004. Correlational tests also showed that across all times L1 literacy measures showed positive and significant relationships with FL reading aloud outcomes (FLRA - Table 14).

Table 14

Relationships between L1 literacy-related outcomes and FL reading aloud at all test times

| | Mid-Test FLRA (fam-unfam) | | Post-Test FLRA (fam-unfam) | | Delayed Post-Test FLRA (fam-unfam) | |
	r^s	sig	r^s	sig	r^s	sig
L1 reading age NGRT	.583**	.003	.513**	.012	.568**	.005
L1 spelling age SWST	.693**	<.0001	.645**	.001	.701**	<.0001
NC Reading Levels	.691**	<.0001	.583**	.004	.767**	<.0001
NC Writing	.516**	.012	.614**	.002	.713**	<.0001

Whilst the influence of L1 reading age on FL reading aloud remained significant at all test times, it diminished slightly after the teaching programme, evidenced by slightly lower r^s values. Interestingly, L1 spelling and teacher-assessed NC levels show highest r^s values, and therefore strongest relationships, at delayed post-test although with the exception of NC writing, these did not result in different values of statistical significance. Spelling age is the most influential, continuously showing r^s values above .6 (strong).

3.3.2. Global FL scores (literacy and general proficiency)

Section 3.3.2 turns to examine FL global performances (collected from additional, combined FL measures) to establish whether FL word card reading aloud scores at each test time were related to underlying FL proficiency. Firstly, global literacy and general proficiency scores at each test time were examined to determine whether FL learning occurred during the teaching cycle and if any attrition occurred between post- and delayed post-test (Tables 15-17). These composite scores were derived from the following

datasets: Global FL literacy comprised FL reading aloud (target-like production of specific sounds/word) in sentences plus FL reading comprehension (reading for global meaning and detail). It should be recalled that the reading aloud task was different from the FL reading aloud word card task reported in RQ1 and RQ2. FL general proficiency was derived from FL receptive vocabulary (recognising words and linking to pictorial representations) plus FL elicited imitation (sentence reproduction).

At pre-test, FL global literacy and general proficiency scores showed wide ranges in achievement across the sample with scores for spoken language measures ranging between 15-57 marks out of a total of 116 (receptive vocabulary, N = 29; elicited imitation, N = 87). For FL literacy, scores were similarly diverse with children's performances ranging from 1 to 9.5 out of a total of 18 (reading aloud, N = 10; reading comprehension, N = 8 – Table 15). By post-test, both means and medians had risen, but it should be noted that there was still a substantial range in achievement scores in both FL general proficiency and global FL literacy variables (Table 16). At delayed post-test, both means and medians had risen slightly for FL general proficiency and had dropped slightly for the FL global literacy scores (Table 17). Again, the range in scores remained wide between the higher and lower achievers.

Table 15

Descriptive statistics pre-test FL general proficiency and FL global literacy (N = 23)

| | Pre-test | | | | | | |
| | | | | Range | | Skewness | Kurtosis |
	Mean	Median	SD	Min	Max	(S.E. Skewness)	(S.E. Kurtosis)
FL general proficiency (max 116)	36.609	41.000	9.981	15.0	57.0	-.822 (.481)	.950 (.935)
FL global literacy (max 18)	5.370	5.500	2.2270	1.0	9.5	.247 (.481)	-.250 (.935)

Table 16

Descriptive statistics post-test FL general proficiency and FL global literacy (N = 23)

| | Post-test | | | | | | |
| | | | | Range | | Skewness | Kurtosis |
	Mean	Median	SD	Min	Max	(S.E. Skewness)	(S.E. Kurtosis)
FL general proficiency (max 116)	45.391	46.000	10.4174	21.0	66.0	-.413 (.481)	.346 (.935)
FL global literacy (max 18)	6.957	6.500	2.4117	3.5	13.0	.482 (.481)	.176 (.935)

Table 17

Descriptive statistics delayed post-test FL general proficiency and FL global literacy (N = 23)

| | Post-test | | | | | | |
| | | | | Range | | Skewness | Kurtosis |
	Mean	Median	SD	Min	Max	(S.E. Skewness)	(S.E. Kurtosis)
FL general proficiency (max 116)	46.652	48.000	10.7602	26.0	67.0	-.304 (.481)	-.379 (.481)
FL global literacy (max 18)	6.500	6.500	2.5628	2.5	12.5	.433 (.481)	-.130 (.935)

Learners' progress on FL general proficiency and FL literacy constructs will now be explored to illustrate broader FL learning during the overall teaching programme (Tables 18 and 19). As these tests also comprised two literacy-related test items, it is interesting to note progress made beyond the sub-skills of FL phonological decoding measured through the FL reading aloud word card test. These analyses also explore whether this broader FL learning was durable by examining scores between post- and delayed post-test.

Table 18 shows that from pre-test to post-test (immediately after the teaching programme), students made statistically significant progress across both FL combined measures. Global general proficiency scores

(pre-test mdn = 39.609; post-test mdn = 45.391, z = 2.930, p = 0.56; r = 0.43) and global literacy scores increased moderately (literacy: pre-test mdn = 5.50; post-test mdn = 6.50; z = 2.930, p = .003; r = 0.43). Between post-test and delayed post-test, medians either remained the same or increased slightly but were not statistically significant (Table 19), showing that the FL learning was relatively durable.

Table 18
Pre-Test to Post-Test FL general proficiency and FL global literacy scores – Wilcoxon Signed Rank Tests (N = 23)

	Median		z	p	S.E.	Effect Size (Cohen's *d*)
	Pre-test	Post-test				
FL General Proficiency	39.609	45.391	3.771	<.0001	30.763	0.56
FL Global Literacy	5.50	6.50	2.930	.003	28.666	0.43

Note. Max score FL general proficiency 116; FL global literacy 18.

Table 19
Post-Test to Delayed Post-Test FL general proficiency and FL global literacy scores – Wilcoxon Signed Rank Tests (N = 23)

	Median		z	p	S.E.	Effect Size (Cohen's *d*)
	Pre-test	Post-test				
FL General Proficiency	45.391	46.652	1.123	.261	30.708	0.17
FL Global Literacy	6.50	6.50	-1.051	.293	26.646	-0.16

Note. Max score FL general proficiency 116; FL global literacy 18.

In Table 19, comparison of FL global literacy scores using a Wilcoxon Signed Rank Test shows that there was no significant difference in scores between post-test (Mdn = 6.50) and delayed post-test (Mdn = 6.50), z = -1.051, p = .293, thereby showing that the progression made in FL literacy during the teaching intervention appeared relatively long-term. However, general proficiency scores increased slightly (though not significantly) between post-test (Mdn = 45.391) and delayed post-test (Mdn = 46.652), z = 1.123, p = .261, small effect size r = 0.17). Section 3.3.3 will turn to global FL proficiencies (general proficiency and global literacy) to establish whether these abilities were related to the FL word card reading aloud scores at each test time.

3.3.3. Global FL and FL reading aloud scores

The relationships between FL general proficiency and FL global literacy with FL reading aloud (word card) outcomes at all test times were explored using Spearman's Rho correlations (Table 20). All pre-test FL outcomes (with the exception of FL global literacy at mid-test FLRA) showed statistically significant relationships with word reading outcomes at mid-, post-, and delayed post-test times. The influence of FL outcomes increased in strength immediately after the teaching programme (at post-test) and also at delayed post-test. General proficiency and literacy-based FL outcomes showed, broadly speaking, the same strength of influence at post-test, evidenced by similar r^s values.

Table 20
Relationships between combined FL outcomes and FL reading aloud word card (familiar and unfamiliar) at all test times

	Mid-Test FLRA (fam-unfam)		Post-Test FLRA (fam-unfam)		Delayed Post-Test FLRA (fam-unfam)	
	r^s	sig	r^s	sig	r^s	sig
Pre-test FL General Proficiency	.480*	.020	.486*	.019	.568**	.005
Pre-test FL Global Literacy	.379	.074	.561**	.005	.570*	.005
Post-test FL General Proficiency	.643**	.001	.552**	.006	.626**	.001
Post-test FL Global Literacy	.612**	.002	.605**	.002	.690**	<.0001
Delayed Post-test FL General proficiency	.501*	.015	.535**	.009	.562**	.005
Delayed Post-test FL Global Literacy	.583**	.004	.786**	<.0001	.781**	<.0001

Whilst it seems likely from extant literature that these FL and L1 variables could be predictors of FL outcomes including FL reading aloud outcomes, the sample size (n = 23) precluded statistical analysis to identify such predictors (Field, 2014).

4. Discussion

4.1. To what extent do learners engaged in systematic and explicit FL phonics instruction make progress in reading aloud familiar and unfamiliar FL words at mid-, post- and delayed post-test?

For familiar word card reading aloud, the children made slow but statistically significant progress both during and immediately after the FL phonics teaching programme. The slow progression is evidenced by a low mean 2.174/median 2.00 at post-test out of a total score of 5. Unfamiliar words and application of sound/spelling links showed similar patterns. Ten weeks into the phonics instruction programme this sub-sample (n = 23) had made almost no progress in applying learned FL sound/spelling links to novel words. The children's progression (for unfamiliar word cards) between mid- and post-test was slow (mean .9130/median 1.00 at post-test) but statistically significant with a moderate effect size (r^s = .40), likely due to a zero-starting point. These findings resonate with prior research in a similar context which found that sound/spelling links, particularly those taught in French (an irregular orthography) developed slowly and that children's reproductions tended to reflect L1 sound/spelling links (Cable et al., 2010). An investigation of phonics instruction conducted in English secondary schools (children aged 11-13) found that, after 16 weeks of systematic and targeted weekly sessions, some participants still failed to successfully decode written FL non-words through a sound-alike test (Woore et al., 2018). It seems likely that development of FL reading aloud skills in both familiar and unfamiliar contexts is a long-term pedagogic endeavour and, consequently, progress may have been further hindered by the relatively low amounts of weekly instructional time, which is known to affect primary FL outcomes (Graham et al., 2017).

Like other primary and secondary school studies (Cable et al., 2010; Woore et al., 2018), this study found that the range of performances was diverse at all test times. Almost one third of the sample (n = 7) scored zero for reading familiar words at mid-test, ten weeks after the phonics instruction commenced and, after 23 weeks of learning, there were still children who were unable to recognise and pronounce familiar words and to decode unfamiliar words. Correlational analyses showed that during and after the teaching programme both L1 reading proficiency and FL proficiencies were related to achievement at FL reading aloud. It could be possible that the learners who apparently made no progress following the phonics instruction programme had lower L1 literacy skills. It should also be noted that, in this sample, L1 spelling ages showed considerable range (7.08 years to 17.0 years). This measure was found to have strong relationships with FL outcomes in this study and L1 spelling is also known to be influenced by L1 phonics instruction (Castles et al., 2018), therefore L1 spelling ranges could have explained some of the range in FL reading aloud scores. Perhaps the phonics pedagogy should have focused more systematically on practising writing both exemplar and other words incorporating the targeted sound/spelling links. On a broader scale, it should be noted that diversity in FL outcomes is a known phenomenon in FL education for young learners. Other studies across Europe have found that FL oracy and literacy scores vary considerably amongst learners of the same ages; even within the same classes (Cable et al., 2010; Enever et al. 2011; Graham et al., 2017).

Finally, the kinds of sound/spelling links taught also appeared to matter to some extent. The best performances at post-test for familiar words centred around novel FL sounds (e.g., nasal vowels) which had to be linked to existing L1 spellings. Those involving extant L1 sounds which required new mappings to different FL spellings seemed more problematic. This perhaps could be explained by the distinctiveness of novel FL sounds which are easier to perceive and therefore link and/or remap to their respective spellings (Flege, 1995). It should also be noted that the phonics instruction which involved both gestures and pictures worked particularly well as clues for the articulatory properties of nasality. For example the nasal vowel /ɔ̃/ was linked to the exemplar word coch<ON> (*pig*). This was accompanied by a picture of a pig and whole class enactment of an action which involved pushing the tip of the nose to resemble a snout whilst saying the sound. This, of course, drew attention to the nasality and potentially acted as a retrieval cue for articulatory properties. However, no phonetic analysis of sound perception or production was involved in the study, therefore this explanation should be treated with caution.

4.2. Are performances on reading aloud familiar and unfamiliar word cards different at mid-, post- and delayed post-test?

Similar patterns for familiar and unfamiliar words showed that learners made slow but significant progress at reading both word types during the teaching. Comparison of familiar versus unfamiliar word card performance at each test time also showed consistently significant differences. In other words, just 10 weeks after the phonics programme started, familiar words were more successfully decoded than unfamiliar words (p = .001; r^s = 0.49 moderate). This preference was replicated at post-test (p < .0001; r^s = -.057) and delayed post-test (p < .0001; r^s = -.052). At each time point learners were significantly better at reading familiar words compared to unfamiliar words.

These results show that enhanced practice and exposure is key in phonics instruction in developing reading aloud of familiar and unfamiliar words. It demonstrates the importance of familiarity with words (item-based learning) to ensure accurate decoding, in line with L1 literature (Share, 1999) and frequency of presentation (Share, 1999). This could also be linked to active decoding and the development of better quality lexical representations (spelling, sound and meaning) in memory (Castles et al., 2018; Ehri, 2014). However, there are a number of complex issues surrounding the difference in performance between known and unknown words. Perhaps the multimodal (enhanced) nature of presentation of phonics exemplar words (written words, pictures, gestures plus meaning) could have boosted the memorisation of familiar words such as *s<IN>ge* (monkey) (Boyer & Ehri, 2011).

Due to frequency of presentation and rehearsal, it is possible that these exemplar words were not actually phonologically decoded in a conscious manner and could have been linked as whole word forms to their corresponding phonological representation. It is also feasible that the children had memorised the oral form of the word and could have generated this from the picture on the word card but it is, nonetheless interesting, that when the visual was presented alongside the written form, some learners were still unable to suppress the corresponding L1 mappings and produced a non-target-like form for the familiar word after 23 weeks of phonics instruction. Again, this seems to suggest that learning FL sound/spelling links takes time and plenty of practise. Historic pedagogic advice and anecdotal teacher opinion contends that FL reading should be avoided until reliable phonological representations of words have been established so that these will counteract L1-based mispronunciations (Hurrell, 1999). This study instead suggests that print-induced L1 mispronunciations will result regardless of timing for L1 literate learners. It seems possible that increasing success rates at post-test show that early opportunities to tackle FL reading can yield eventual target-like decoding. However, as section 5.1 notes, it is beyond the remit of this study to attribute any outcomes directly to the efficacy of FL phonics instruction and pedagogic recommendations would benefit from further research.

Evidently the phonological decoding of unfamiliar words is the bona fide indicator of application of conscious sound/spelling link knowledge. However, particular sound/spelling links which were more successfully realised in familiar forms performed better in unfamiliar contexts too, and this seems to indicate that there could be a link between memorisation of exemplars and conscious phonological decoding. Whilst not a focus of analysis, learners tended to make L1 related errors when phonological decoding such as substituting ON = /ɒn/ for ON = /ɔ̃/. This seems to show that a degree of automatized L1 phonological decoding was indeed activated, even for conscious or effortful decoding of unfamiliar words and successful performance could have indicated whether children were able to override these automatized L1 sound/spelling link knowledge. It appears that, whilst many of them were making progress in this regard, reading in an additional language is a dual (or multiple) language affair, and overriding competing L1 knowledge is challenging (Cable et al., 2010; Grabe & Stoller, 2002; Koda, 2008; Woore, 2009). Perhaps the learners could have benefited from more opportunities to apply practised sound/spelling links in novel contexts, thereby strengthening recall of FL sound/spelling link knowledge. The phonics instruction programme could have included more opportunities to practise reading a range of unknown words.

4.3. To what extent do L1 literacy measures and FL global proficiency scores relate to FL reading aloud proficiency at each test time?

Statistical analyses showed that all L1 literacy measures (both standardised test scores and teacher assessments) correlated strongly and positively at all test times with FL word reading aloud test scores. This indicated, as expected from the literature, that successful L1 readers and writers were likely to be more successful at FL reading aloud. Wider research has found that young learners FL literacy and indeed FL

outcomes generally are likely to be influenced by L1 literacy (Courtney et al., 2017; Sparks & Ganschow, 2012). The relationships between L1 literacy and FL word card reading aloud decreased very slightly at post-test, then increased again at delayed post-test. This was despite the differences between post-test scores and delayed post-test scores not being statistically significant. It seems possible that this could relate to a particularly low start for FL reading aloud scores which increased immediately following the FL phonics instruction programme. It could also show that some of the lower achieving L1 literacy learners increased their FL reading aloud skills as their L1 literacy skills were improving too. This was not evidenced as the L1 data collection was focused on the beginning of the school year, however, it is likely that L1 literacy skills would be boosted markedly during the last two years of primary school when pupils are preparing for national standardised literacy and numeracy tests.

Both L1 reading age and L1 spelling age were influential at all test times, although spelling age was marginally more important, tending to demonstrate strong rather than moderate relationships. This could be because the L1 reading age test examines word recognition and reading comprehension by asking student to select the correct written word from a range of options to complete a gapped sentence, whilst the SWST spelling test involves typing or writing a target word in a gapped sentence from an auditory cue (e.g. orthographic decoding). Arguably the SWST would act as a more direct indicator of L1 relationships between sounds and letters. It could also be relevant to note that the SWST scores showed greater range than the NGRT scores and this perhaps influenced the statistical analyses. This could also reflect links with phonics knowledge and spelling evidenced in the L1 literature (Rosenthal & Ehri, 2011).

Interestingly, FL outcomes were generally related to FL reading aloud word card scores at all test times, with the exception of FL global literacy scores at mid-test. So, if learners were high achievers at either FL literacy or FL global proficiency, then they tended to score higher for FL word card reading aloud. As discussed in the results section, the sample size was not large enough to explore these variables as predictors, so it is impossible to discern whether these FL proficiencies are likely to have determined performance. It is also important to note that relationships between FL outcomes and FL reading aloud could have been mediated by other factors such as learner self-efficacy and motivation, as earlier young learner FL studies have noted links between attitudinal factors and learning outcomes (Courtney, 2017; Graham et al., 2017).

5. Conclusion

To summarise, over the course of a 23-week programme of systematic weekly phonics activities where French sound/spelling links were taught and practised, learners made steady progress. The reading aloud process showed great variation inter-participant, but the overall learning, whilst slow, tended to be durable and was evident at least seven weeks after the instructional programme finished. The influence of L1 literacy-related skills, particularly L1 spelling age, was important for FL reading outcomes. In other words, skilled L1 readers had a clear advantage for reading aloud FL words. For all learners, familiar FL phonics exemplar words were more accurately read aloud than unfamiliar words, but after 23 weeks of instruction, progress applying learned sound/spelling links to novel (unfamiliar) contexts started to emerge. This study shows that learning to read aloud FL words is a slow process but that children do make progress. On a broader scale, this study confirmed that young, beginner FL learners are able to make meaningful progress in limited input classroom settings. It was also apparent that whilst L1 literacy skills have facilitative role in FL outcomes, the process of re-mapping (and/or suppressing) L1 sound/spelling links is a complex and arduous one.

5.1. Limitations

It is important to recall that this was not a direct evaluation of the phonics instructional programme alone, as the children were engaged in a rich programme of oracy and literacy activities, which has been documented in the supplementary materials. This is in line with wider literacy-related research which evidences the importance of both phonics instruction and meaning-based text activities to support learning (Adams, 1990). To evaluate the direct implications of the phonics instruction package, one school would have had to have only the phonics instructional programme. As little is known about best practice for FL phonics instruction or its effectiveness, this study undertook to investigate the possible outcomes and learning processes in a limited input, limited timeframe FL setting. It is also important to consider that teacher factors such as linguistic ability, language specialist training, and teaching time are likely to underpin FL outcomes (Graham et al., 2017) and could have influenced progression in this context. Arguably, the teacher/researcher

stance meant that this study was implemented by someone who was heavily invested in the teaching/learning process.

5.2. A future research agenda

As it cannot be assumed that the benefits for phonics in English L1 literacy learning can simply be transferred to any FL context, future research should aim to investigate the effectiveness of FL phonics in primary schools in a controlled experiment. It would also be useful to examine the learning of sound/spelling links of different foreign languages in young learner instructed settings as well as the teaching and learning of FL literacy for learners with different L1 (non-alphabetic) backgrounds such as Chinese.

Funding

Acknowledgements

References

Adams, Marilyn J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, Massachusetts: MIT Press.

Antony, Jason L, Lonigan, Christopher J., Burgess, Stephen R., Driscoll, Kimberley, Phillips, Beth M., & Cantor, Brenlee G. (2002). Structure of pre-school phonological sensitivity: Overlapping sensitivity to rhyme, words, syllables and phonemes. *Journal of Experimental Child Psychology*, *85*, 65-92.

Apel, Kenn, Wilson-Fowler, Elizabeth B., Brimo, Danielle, & Perrin, Nancy A. (2012). Metalinguistic contributions to reading and spelling in second and third grade students. *Read Write*, *25*, 1283-1305.

Bassetti, Bene, & Atkinson, Nathan. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, *36*(1), 67-91.

Biemiller, Andrew, & Slonim, Naomi (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, *93*(3), 498-520.

Boyer, Nancy, & Ehri, Linnea (2011). Contribution of phonemic segmentation instruction with letters and articulation pictures to word reading and spelling in beginners. *Scientific Studies of Reading*, *15*(5), 440-470.

Cable, Carrie, Driscoll, Patricia, Mitchell, Rosamond, Sing, Sue, Cremin, Teresa, Earl, Justine, Eyres, Ian, Holmes, Bernardette, Martin, Cynthia, & Heins, Barbara. (2010). *Primary modern languages: A longitudinal study of language learning at Key Stage 2*. Nottingham, United Kingdom: Department for Children, Schools & Families.

Campfield, Dorotea (2015). Lexical difficulty – using elicited imitation to study child L2. *Language Testing*, *34*(2), 1-12.

Castles, Anne, Rastle, Kathleen, & Nation, Kate (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*, 5-51.

Chall, Jeanne (1967). *Learning to read: The great debate.* New York, New York: McGraw Hill.

Charlton, Adeline (2014). *The second language acquisition of French nasal vowels by advanced English learners*. Second Language Acquisition of Phonology 2012 workshop, University of York, 2012. Retrieved from http://www-users.york.ac.uk/~sh581/POSTER-ELOL%20Charlton.pdf.

Courtney, Louise (2017). Transition in modern foreign languages: a longitudinal study of motivation for language learning and second language proficiency. *Oxford Review of Education*, *43*(4), 462-481.

Courtney, Louise, Graham, Suzanne, Marinis, Theo, & Tonkyn, Allan (2017). Individual differences in early language earning: A study of English learners of French. *Applied Linguistics, 38*(6), 824-847.

Cummins, Jim (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*, 222–251.

Deacon, S. Hélène, & Kirby, John R. (2004). Morphological awareness: Just more phonological? The roles of morphological awareness and phonological awareness in reading development. *Applied Psycholinguistics*, *25*, 223-238.

Department of Education, Science & Training (2005). *Teaching reading: Report and recommendations.* National Inquiry into the Teaching of Literacy.

DfE (2013). *The national curriculum in England. Key Stages 1 and 2 Framework Document.* Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/425601/PRIMARY_national_curriculum.pdf.

DfES (2005). *The key stage 2 framework for languages*. Nottingham, United Kingdom: DfES Publications.

Dunn, Lloyd M., & Dunn, Leota M. (2007). *PPVT-4 Pearson Picture Vocabulary Test*. Minneapolis, Minnesota: Pearson Assessments.

Dunn, Lloyd M., Dunn, Leota M., Whetton, Chris, & Burley, Juliet (1997). *British Picture Vocabulary Scale*. Windsor, United Kingdom: NFER.

Durgunoğlu, Aydin Y., Nagy, William B., & Hancin-Bhatt, Barbara J. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology*, *85*, 453-465.

Ehri, Linnea, C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory and vocabulary learning. *Scientific Studies of Reading*, *18*(1), 5-21.

Ehri, Linnea, C. (1999). Phases of development in learning to read words. In Jane Oakhill & Roger Beard (Eds.), *Reading development and the teaching of reading* (pp. 79-108). Oxford, United Kingdom: Blackwell.

Ehri, Linnea, C. (1992). Reconceptualising the development of sight word reading and its relationship to recoding. In Phillip B. Gough, Linnea C. Ehri, & Rebeca Treiman (Eds.), *Reading acquisition* (pp. 107-143). Hillsdale, New Jersey: Lawrence Erlbaum.

Ehri, Linnea C., Nunes, Simone R., Stahl, Steven A., & Willows, Dale M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Journal of Direct Instruction*, *2*(2), 121-166.

Enever, Janet, Krikhaar, Evelien, Lindgren, Eva, Lundberg, Gun, Mihaljević Djigunović, Jelena, & Muñoz, Carmen (2011). *ELLiE – Early Language Learning in Europe.* London, United Kingdom: British Council.

Erlam, Rosemary (2006). Elicited imitation as a measure of L2 implicit knowledge: an empirical validation study. *Applied Linguistics*, *27*, 464-491.

Escudero, Paola, & Boersma, Paul (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551–585.

Escudero, Paola & Wanrooij, Karin (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech*, *53*(3), 343-365.

Flege, Jim E. (1995). Second language speech learning. Theory, findings, and problems. In Winifred Strange (Ed.), *Speech perception and linguistic experience. Issues in cross-language research* (pp. 233-277). Baltimore, Maryland: York Press.

Garton, Susan, Copland, Fiona, & Burns, Ann (2011). *Investigating global practices in teaching young learners*. London, United Kingdom: British Council.

Genesee, Fred, Geva, Esther, Dressler, Cheryl, & Kamil, Michael L. (2008). Cross-linguistic relationships in second language learners. In Diane A. August & Timothy Shanahan (Eds.), *Developing reading and writing in second language learners* (pp. 63-93). New York/Abingdon: Routledge.

GL Assessment. (2018). *New Group Reading Test* (NGRT). Retrieved from https://www.gl-assessment.co.uk/products/new-group-reading-test-ngrt/).

GL Assessment. (2018). *Single Word Spelling Test* (SWST). Retrieved from https://www.gl-assessment.co.uk/products/single-word-spelling-test-swst/).

Goswami, Usha (2006). Orthography, phonology and reading development: A cross-linguistic perspective. In R. Malatesha Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 463-480). Mahwah, New Jersey: Laurence Erlbaum.

Gough Phillip B., & Wren, Sebastian (1999). Constructing meaning: The role of decoding. In Jane Oakhill & Roger Beard (Eds.), *Reading development and the teaching of reading* (pp. 59-78). Oxford, United Kingdom: Blackwell.

Grabe, William, & Stoller, Frederika L. (2002). *Teaching and researching reading.* Harlow, Essex: Pearson Education.

Graham, Suzanne, Courtney, Louise, Marinis, Theo, & Tonkyn, Allan (2017). Early language learning: The impact of teaching and teacher factors. *Language Learning*, *67*(4), 1-37.

Hatcher, Peter J., Hulme, Charles, & Ellis, Andrew W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, *65*(1), 41-57.

Henbest, Victoria S., & Apel, Kenn (2017). Effective word reading instruction: What does the evidence tell us? *Communication Disorders Quarterly*, *39*(1), 303-311.

Hurrell, Alison (1999). The four language skills. In Patricia Driscoll & David Frost (Eds.), *The teaching of modern foreign languages in primary school* (pp. 67-87). London, United Kingdom: Routledge.

Kaefer, Tanya (2016). Integrating orthographic and phonological knowledge in early readers: Implicit and explicit knowledge. *Child Development Research*. https://doi.org/10.1155/2016/6036129.

Koda, Keiko (1992).The effects of lower-level processing skills on FL reading performance: Implications for instruction. *Modern Language Journal, 76*(4), 502-512.

Koda, Keiko (2007). Reading and language learning: Cross-linguistic constraints on second-language reading development. *Language Learning, 57*(s1), 1-44.

Koda, Keiko (2008). Impacts of prior literacy experience on second-language learning to read. In Keiko Koda & Annette Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first and second-language literacy development* (pp. 68-96). New York/Abingdon: Routledge.

Koutsompou, Violetta-Irene (2015). The use of literature in the language classroom. Methods and Aims. *International Journal of Information and Education Technology*, *5*(1), 74-79.

Lado, Robert (1957). *Linguistics across cultures: Applied linguistics and language teachers.* Ann Arbor, Michigan: University of Michigan Press.

Landerl, Karin. (2006). Reading acquisition in different orthographies: Evidence from direct comparison. In R. Malatesha Joshi & P. G. Aaron (Eds.), *The handbook of orthography and literacy* (pp. 513-530). Mahwah, New Jersey: Lawrence Erlbaum.

Lesaux, Nonie K., Geva, Esther, Koda, Keiko, Siegel, Linda S., & Shanahan, Timothy (2008). Development of literacy in second language learners. In Diane A. August & Timothy Shanahan (Eds.), *Developing reading and writing in second language learner* (pp. 27-60). New York/Abingdon: Routledge.

Lloyd, Janet, & Wernham, Susan (1992). *Jolly phonics: A handbook for teaching reading, writing and spelling.* Essex, United Kingdom: Jolly Learning.

Lonsdale, Deryle, Graham, C. Ray, Kennington, Casey, Johnson, Aaron, & McGhee, Jeremiah (2008). Scoring an oral language test using automated speech recognition. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, & Daniel Tapias (Eds.), *Proceedings of the 6th International Conference on Language Resources and Evaluation* (pp. 1604-1610). Marrakech, Morocco: European Language Resources Association.

Macaro, Ernesto, & Erler, Lynne (2008). Decoding ability in French as a foreign language and language learning motivation. *The Modern Language Journal*, *95*(4), 496-517.

Macaro, Ernesto, & Mutton, Trevor (2009). Developing reading achievement in primary learners of French: Inferencing strategies versus exposure to graded readers. *Language Learning Journal*, *37*, 165-182.

Martinez, Márquez (2016). *The acquisition of French nasal vowels: From first language allophony to second language phonological contrast.* Unpublished PhD dissertation. Indiana University.

McBride Chang, Catherine, & Ho, Connie Suk-Han (2005). Predictors of beginning reading in Chinese and English: A two year longitudinal study of Chinese kindergartners. *Scientific Studies of Reading*, *9*(2), 117-144.

McKay, Penny (2006). *Assessing Young Learners.* Cambridge, United Kingdom: Cambridge University Press.

Miskin, Ruth (2011). *Read Write Inc phonics handbook*. Oxford, United Kingdom: Oxford University Press.

Mitchell, Rosamond, Myles, Florence, & Marsden, Emma (2019). *Second language learning theories*. London, United Kingdom: Hodder Arnold.

Murphy, Victoria, Macaro, Ernesto, Alba, Sonia, & Cipolla, Claudia (2015). The influence of learning a second language in primary school on developing first language literacy skills. *Applied Psycholinguistics*, *36*(5), 1133-1153.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research Literature on reading and its implications for reading instruction*. Washington, District of Columbia: National Reading Panel.

Neveu, Frank, Muni-Toké, Valelia, Klinger, Thomas, Durand, Jacques, Mondada, Lorenza, & Prévost, Sophie (2010). Evaluation des voyelles nasales en français L2 en production: de la nécessité d'un corpus multi-tâches. In Franck Neveu, Valelia Muni Toke, Thomas Klingler, Jacques Durand, Lorenz Mondada, & Sophie Prévost (Eds.), *Proceedings of the Congrès Mondial de Linguistique Française*. Paris, France: EDP Sciences. https://doi.org/10.1051/cmlf/2010119.

Nimz, Katharina (2018). *Sound perception and production in a foreign language: Does orthography matter?* Potsdam, Germany: Universitätsverlag Potsdam.

Oakhill, Jane, & Beard, Roger (1999). *Reading development and the teaching of reading*. Oxford, United Kingdom: Blackwell Publishing.

Perfetti, Charles, & Stafura, Joseph (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(1), 22-37.

Porter, Alison (2014). *An early start to French literacy: Learning the spoken and written word simultaneously in English primary schools.* Unpublished PhD Thesis. Southampton, United Kingdom: University of Southampton.

Ricketts, Jessie, Bishop, Dorothy V., Pimperton, Hannah, & Nation, Kate (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading*, *15*(1), 47-70.

Rosenthal, Julie, & Ehri, Linnea, C. (2011). Pronouncing new words aloud during the silent reading of text enhances fifth graders' memory for vocabulary words and their spellings. *Read Write*, *24*, 921-950.

Share, David L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*, 95-129.

Shmidman, Adina, & Ehri, Linnea, C. (2010). Embedded picture mnemonics to learn letters. *Scientific Studies of Reading, 14*(2), 159-182.

Sparks, Richard L. & Ganschow, Leonore (1991). Foreign language learning differences: Affective or native language aptitude differences. *The Modern Language Journal*, *75*(1), 3-16.

Sparks, Richard L., Patton, Jon, Ganschow, Leonore, & Humbach, Nancy (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Language Learning*, *59*(1), 203-243.

Sparks, Richard L., Patton, Jon, & Ganschow, Leonore (2012). Profiles of more and less successful L2 learners: A cluster analysis study. *Learning and Individual Differences*, *22*(4), 463-472.

Stanovich, Keith E., & Stanovich, Paula, J. (1999). How research might inform the debate about early reading acquisition. In Jane Oakhill & Roger Beard (Eds.), *Reading development and the teaching of reading* (pp. 12-41). Oxford, United Kingdom: Blackwell Publishing.

Stuart, Morag, Masterson, Jackie, & Dixon, Maureen. (1999). Learning to read words turns listeners into readers. How children accomplish this transition. In Jane Oakhill & Roger Beard (Eds.), *Reading development and the teaching of reading* (pp. 109-130). Oxford, United Kingdom: Blackwell Publishing.

Rose, Jim (2006). *Independent review of the teaching of early reading*. Nottingham, United Kingdom: Department of Education and Skills.

Torgerson, Carole J., Brooks, Greg, & Hall, Jill (2006). *A systematic review of the research literature on the use of phonics in the teaching of reading and spelling*. DfES Research Report RR711.

Tyler, Michael D., Best, Catherine T., Faber, Alice, & Levitt, Andrea G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica, 71*(1), 4–21.

Vellutino, Frank R. & Fletcher, Jack M. (2005). Development dyslexia. In Margaret J. Snowling & Charles Hulme (Eds.), *The science of reading: A handbook* (pp. 362-378)*. Oxford, United Kingdom: Blackwell.

Verhoeven, Ludo (2011). Second language reading. In Michael L. Kamil, David, Pearson, Elizabeth B. Moje, & Peter Afflerbach (Eds.), *Handbook of reading research* (pp. 661-683). New York/London: Routledge.

Woore, Robert (2009). Beginners' progress in decoding L2 French: Some longitudinal evidence from English modern foreign languages classrooms. *Language Learning Journal, 37*(1), 3-18.

Woore, Robert (2011). *Investigating and developing beginner learners' decoding proficiency in second language French: An evaluation of two programmes of instruction*. Unpublished PhD thesis*. Oxford: University of Oxford.

Woore, Robert, Graham, Suzanne, Porter, Alison, Courtney, Louise & Savory, Claire (2018). *Foreign language education unlocking reading (FLEUR): A study into the teaching of reading to beginner learners of French in secondary school.* Oxford, United Kingdom: University of Oxford.

Wyse, Dominic & Styles, Morag (2007). Synthetic phonics and the teaching of reading: The debate surrounding England's Rose Report. *Literacy, 41*(1), 37-42.

Alison Porter, University of Southampton
A.Porter@soton.ac.uk

EN | **Alison Porter** is a lecturer in Applied Linguistics at the University of Southampton. Her primary research area is young language learning in instructed settings. She is particularly interested in the teaching and learning of foreign language literacy and has participated in teaching and learning research in both early (ages 5-7) and late (ages 9-11) primary school classrooms. She also taught foreign languages in English primary schools for ten years. Most recently, Dr. Porter collaborated on a large-scale teaching intervention study in early secondary school classrooms (ages 11-13) which aimed to evaluate the effects of different types of reading instruction: phonics and strategy use. In 2017, Dr. Porter established a Southampton University Primary Languages (SUPL) Research and Practice partnership with Emeritus Professor Ros Mitchell. SUPL currently collaborates with 20 local primary schools to co-construct, implement, and evaluate foreign language literacy practices in classrooms. Participating teachers are concurrently trained in classroom research methods, and several small-scale investigations have already resulted. Dr. Porter is also a member of the Research in Primary Language network (RiPL), founded by the University of Essex, which engages with practitioners on a local level as well as national stakeholders in a bid to inform policy and sustainability issues relating to primary foreign language provision.

ES | **Alison Porter** es una profesora de Lingüística Aplicada en la Universidad de Southampton. Su área de investigación es principalmente el aprendizaje de la lengua en niños en un contexto educativo. Se ocupa sobretodo de enseñanza y aprendizaje de alfabetización de una lengua extranjera y ha contribuido en investigaciones sobre la enseñanza y el aprendizaje tanto en niños de 5-7 años como en los de 9-11 años. Impartió clases de lenguas extranjeras en escuelas primarias durante diez años. Más recientemente, la profesora Porter ha colaborado a un estudio a gran escala sobre la intervención de la enseñanza en las escuelas secundaria (11-13 años de edad) que tenía como objetivo evaluar los efectos producidos por diferentes tipos de enseñanza de lectura: instrucción fónica y uso estratégico. En 2017, la profesora Porter estableció la asociación Southampton University Primary Languages (SUPL) Research and Practice junto con el Profesor émerito Ros Mitchell. Hoy en día SUPL colabora con 20 escuelas primarias locales para desarrollar juntos, implementar y evaluar las practicas de alfabetización de lenguas extranjeras en clase. Los docentes que participan son formados al mismo tiempo en métodos de investigación desarrollando ya numerosas investigaciones a pequeña escala. La profesora Porter es también miembro de Research in Primary Language network (RiPL), fundado por la University of Essex, que se relaciona con tanto con profesionales a nivel local como con stakeholders a nivel nacional con el intento de informar sobre cuestiones de política y sostenibilidad relacionadas con la oferta de la lengua extranjera en la escuela primaria.

IT | **Alison Porter** è Ricercatrice in Linguistica applicata presso l'Università di Southampton. La sua area di ricerca principale è l'apprendimento linguistico dei bambini in contesto di istruzione formale. È particolarmente interessata all'insegnamento e apprendimento della letto-scrittura nella lingua straniera e ha partecipato a ricerche sull'insegnamento e apprendimento in classi iniziali (5-7 anni) ed avanzate (9-11) nella scuola primaria. Ha inoltre insegnato lingua straniera nella scuola primaria inglese per dieci anni. Ultimamente, Porter ha collaborato a un ampio studio sull'intervento didattico in classi iniziali della scuola secondaria (11-13 anni) finalizzato alla valutazione degli effetti di tipi diversi di insegnamento della lettura: il metodo fonetico e l'uso di strategie. Nel 2017, Porter ha stabilito una collaborazione dell'Università di Southampton con il professore emerito Ros Mitchell per Southampton University Primary Languages (SUPL) Research and Practice (SUPL). Attualmente, SUPL collabora con 20 scuole locali per la creazione, implementazione e valutazione di pratiche di alfabetizzazione per la classe. Gli insegnanti coinvolti vengono al contempo istruiti sui metodi di ricerca per la classe e ne sono già derivati diversi studi su piccola scala. Porter è anche membro della Research in Primary Language network (RiPL) fondata dall'Università dell'Essex, che si confronta con gli operatori del settore a livello locale e con le parti interessate a livello nazionale nel tentativo di orientare questioni di politiche e sostenibilità relative all'insegnamento delle lingue straniere nella scuola primaria.