

## **AWARENESS OF L1 FORM-MEANING MAPPINGS CAN REDUCE CROSSLINGUISTIC EFFECTS IN L2 GRAMMATICAL LEARNING**

Kevin McManus, Pennsylvania State University, USA, [kmcmanus@psu.edu](mailto:kmcmanus@psu.edu)

### **Abstract**

The present study examined the extent to which first language (L1) awareness can benefit second language (L2) grammatical learning of the French *Imparfait*, a crosslinguistically complex target feature. Sixty-nine English-speaking learners of L2 French received different types of explicit information about L2 or L2 + L1 form-meaning mappings. A ‘core’ treatment received by all learners consisted of EI about L2 with L2 comprehension-based practice of French sentences. Two further treatments examined the impact of additional (i) EI about L1 and (ii) comprehension practice of L1 sentences. Results from an online self-paced reading test, offline judgment tests (in reading and listening) and an oral sentence completion test with source of knowledge probes showed that performance immediately after the instruction and then six weeks later improved only for learners exhibiting L1 awareness. These results suggest that awareness of L1 form-meaning mappings can benefit L2 grammar learning of a crosslinguistically complex target feature.

Key words: SLA, Explicit instruction, Language awareness, Grammar, Crosslinguistic contrasts

## Introduction

Understanding the ways in which first language (L1) knowledge can influence second language (L2) learning constitutes a major line of second language acquisition (SLA) research to date (Ellis 2006, MacWhinney, 2012; VanPatten, 2015). Research indicates persistent difficulties when a learner's L1 and L2 express the same meanings in different ways (also called negative influence), but fewer difficulties when L1 and L2 express the same meanings similarly (also called positive influence; Amenós-Pons, Ahern, & Fuentes, 2017; Izquierdo & Collins, 2008; McManus, 2013, 2015), even after many years of exposure and/or instruction (Mitchell, Tracy-Ventura, & McManus, 2017; Ayoun, 2013). The French *Imparfait* (IMP), commonly used to express past habituality and past ongoingness (Hoffman, 1995), is one target feature well-documented to reflect these learning difficulties (Amenós-Ponset al., 2017; Howard, 2005; McManus, 2015). One potential explanation for IMP's L2 learning difficulty is that languages are inconsistent in how express the same meanings (see Comrie, 1976; Smith, 1997). As shown in Table 1, for example, past ongoingness and past habituality can be expressed using the same verbal form in French (*Imparfait*, Hoffman, 1995), but these same meanings are expressed in English using separate verbal forms (*Past Progressive* for past ongoingness, *Simple Past*, *used to*, and *would* for past habituality, Comrie, 1976; Tagliamonte & Lawrence, 2000). Therefore, French and English lack consistency in how they express the same meanings, understood to be an important explanation for English-speaking learners' difficulties in the processing and use of *Imparfait* (Howard, 2005; McManus, 2015; McManus & Marsden, 2017).

<TABLE 1>

However, despite considerable understanding about the ways in which L1-L2 form-meaning mapping differences can influence L2 learning (Ellis, 2006; MacWhinney, 2012), very little research has examined the extent to which this evidence base can be used to improve L2 grammar learning. Intervention research designed to reduce the effects of L1-L2 differences on L2 performance has shown that raising learners' awareness about the nature of L1-L2 differences (using EI about L2 and L1) immediately benefitted written, untimed L2 production (Ammar, Lightbown, & Spada, 2010; Horst, White, & Bell, 2010), whereas raising awareness about L2 properties only (using EI about L2 without EI about L1) did not benefit performance on grammaticality judgment tests (Tolentino & Tokowicz, 2014). Consistent with SLA theories that forefront critical roles for explicit knowledge about language in explaining L2 learning outcomes (e.g., DeKeyser, 1997; Roher-Brackin 2018), these lines of research suggest that providing EI about L1 and L2 helped learners develop awareness about L1-L2 differences (understood to be a source of the learning problem) and that meaning-based practice operating on that awareness led to improved L2 performance. This evidence base about L1 awareness in L2 learning, however, is small and has only examined learning outcomes immediately after instruction using untimed tests allowing for planning and conscious reflection. Research is yet to investigate the extent to which these benefits are detectable using delayed measures of language performance tapping a variety of skills (e.g., comprehension, oral production, real-time processing). Incorporating these design elements into the existing evidence base has the potential to majorly advance knowledge and understanding about the durability of these instructional effects and the types of knowledge representations they may benefit (Rebuschat, 2015).

The current study addressed this gap by examining the extent to which awareness of L1-L2 differences can benefit L2 learning of the French *Imparfait*, a target feature well-documented

to be late-acquired due to functional complexity and L1-L2 form-meaning mapping differences (Ayoun, 2013; Bartning, 2009; McManus, 2013, 2015). To examine this question, English-speaking learners of L2 French received one of three types of explicit instruction that differed in terms of type of explicit information (EI) and type of comprehension practice. First, an ‘L2-only’ group received a core of EI about French (L2) form-meaning mappings for past habituality and past ongoingness with comprehension practice of French sentences. Second, an ‘L2+L1prac’ group received the same core of French EI and French comprehension practice plus additional comprehension practice of English (L1) sentences. Third, an ‘L2+L1’ group received the same core (L2 EI, L2 comprehension practice), plus additional English comprehension practice and EI about English form-meaning mappings for past habituality and past ongoingness. This design examined the roles played by additional EI about L1 and comprehension practice of L1 sentences in L2 grammatical learning to advance understanding about (a) the extent to which explicit instruction (about L2 vs. about L2 + L1) might lead to the creation/development of language awareness and (b) the extent to which evidence of language awareness is detectable in a variety of language performance measures immediately after the instruction (posttest) and then six weeks later (delayed posttest). No previous research has investigated these questions for this target feature. In particular, the study examined the extent to which EI designed to make explicit L2 and L1 form-meaning mappings for past habituality and past ongoingness facilitated L2 learning.

In the sections that follow, first discussed is previous research on the acquisition of IMP in L2 French, followed by a review of existing research addressing L1-L2 differences in L2 learning. The potential learning benefits of explicit knowledge about L1 and L2 are then discussed.

## **L2 learning of French *Imparfait***

Research to date examining the L2 learning of French IMP has repeatedly shown it to be a late-acquired verbal form, arguably due its polyfunctionality, its relative infrequency compared to other past time forms (e.g., *Passé Composé*), and its tendency to be used most often with state verbs (Ayoun, 2013; Bartning, 2009; Kihlstedt, 2015; McManus, 2013, 2015). Although IMP's polyfunctionality is an important explanation for difficulties in its acquisition, its persistent restriction to the state verbs *avoir* ('have') and *être* ('be') represents an active line of inquiry (Ayoun, 2013; Kihlstedt, 2015). This body of research suggests challenges using IMP with non-state verb types (but see Salaberry, 2008; McManus, 2013). Some intervention research has investigated the extent to which particular types of instruction (e.g., feedback types, use of multimedia) addressing this lexical restriction might encourage learners to use IMP with a variety of verb types (e.g., Ayoun, 2001; Izquierdo, 2014). For example, Mifka-Profozic (2015) examined how different types of corrective feedback might extend learners' use of IMP to verbs other than states. Although Pretest-Posttest results showed only limited improvement for IMP, instruction did benefit learners' use of other past time forms (e.g., *Passé Composé*). Such findings have led to claims that IMP's lexical restriction, rote-learned use, and late-acquisition could indicate learners are unaware of how to use this verbal form. In other words, IMP might not be initially productive because learners are unaware of the meanings that this form can express (see also McManus, 2015).

Indeed, research has shown that not all of IMP's meanings are acquired together (Howard, 2005; Kihlstedt, 1998). Some research suggests that the order of acquisition of IMP's meanings is influenced by L1 background: Ongoingness acquired before habituality for English-

speaking learners (Howard, 2005), but habituality acquired before ongoingness for Swedish-speaking learners (Kihlstedt, 2002). Howard (2005) suggests that the grammatical features of English, including the existence of English progressive forms, likely makes ongoingness easier to conceptualize for English speakers, which has the potential to facilitate L2 learning due to easier transfer (Slobin 1973). In this same vein, McManus and Marsden (2019) claim that greater variation in how habituality is expressed in English (e.g., *Simple Past, would, used to*) reduces learners' sensitivity to the concept of habituality, thus impeding its availability for transfer (see also Vanek & Hendriks 2015). This account is consistent with research showing greater difficulties in learning IMP's habitual meaning relative to its ongoing meaning among English-speaking learners (Ayoun, 2013; Howard, 2005; McManus, 2015). Taken together, the nature of L1 form-meaning mappings could play an important role in understanding IMP's acquisition (see also Andersen, 1984; Ayoun, 2013; Salaberry, 2008). To date, no research has sought to improve L2 learners' use of IMP through targeted instruction of L1-L2 differences. The next section reviews how previous research has addressed the role of L1-L2 differences of other target features in instructed L2 learning.

### **Addressing L1-L2 differences in L2 learning**

Psycholinguistic research has repeatedly shown that L1 knowledge can majorly influence L2 grammatical learning in two important ways: languages differ in meanings they map to specific forms (or linguistic cues) (MacWhinney, 2012) and L1 experience influences the forms (or cues) that learners attend to (Ellis, 2006).

An important line of inquiry building on this evidence has examined the extent to which providing EI about the L2 can address the negative influence of L1-based processing strategies

and L1-tuned attention in L2 learning (e.g., Cintrón-Valentín & Ellis 2016; VanPatten, 2015).

For example, to address processing difficulties with English articles, Zhao and MacWhinney (2018) trained Chinese-speaking learners of English to attend to noun-based article cues (e.g., plural, non-countable) for determining article selection. In a similar vein, Ellis and colleagues intentionally manipulated learned attention during L2 learning (Cintrón-Valentín & Ellis, 2016; Ellis & Sagarra, 2010): English and Chinese speakers attend to adverbial cues over verbal cues for processing temporal reference, but not all languages reliably express temporal reference adverbially. Three one-hour treatments were provided to examine the extent to which L1-based processing behaviours could be re-oriented to improve L2 learners' processing of temporal information in L2 Latin (i.e., away from L1-based adverbial processing towards more appropriate L2 processing that uses verb-based cues): linguistic explanations about verb inflection cues, highlighting of verb inflection cues to increase saliency, and processing practice of verbs without adverbials. Eye-tracking results immediately after the training indicated that all conditions increased attention to verbs, but most improvement was found for processing practice on verbs without adverbials.

This body of research indicates that L2 learning can be improved by explicitly drawing learners' attention to L1-L2 differences. However, since, in most cases, training has been intensive and short with a tendency not to include delayed posttests, the durability of this type of training is unclear. Future research with delayed posttesting is needed to assess these claims. The next section reviews the small body of research examining how providing EI about the L1 might also benefit the L2 learning of target features with L1-L2 differences

### **Explicit knowledge about L1-L2 differences in L2 learning**

Research examining the extent to which raising learners' awareness about L1-L2 differences can benefit L2 learning is informed by cognitive theories of L2 learning that attribute critical roles to explicit knowledge/awareness (for reviews, see Tyler, 2012a; Tyler & Ortega, 2018). These accounts propose that L2 learning can be facilitated by instruction that builds understandings of language: the meanings of words, word strings, and how words and word strings fit together. For example, Tyler (2012b) proposes that Russian-speaking learners of L2 English struggle to correctly use the English prepositions 'over' and 'at' because these meanings are expressed in Russian with a single preposition, 'za'. Tyler proposes that instruction could address this learning problem by explicitly raising learners' awareness of the meaning implications of these prepositions in L1 and L2.

In this same vein, to address well-documented learning difficulties associated with questions among French-speaking learners of English (see Spada & Lightbown 1999), Spada, Lightbown, and White (2005) provided explicit information about question formation differences and similarities in French and English. Of particular interest was yes/no questions, which tend to be asked in French using declarative word order with raising intonation (e.g., *Jean est chez lui?*, 'John is at home?'), but in English are asked using interrogative word orders (or subject-verb inversion, e.g., *Is John at home?*). Furthermore, declarative word orders for yes/no questions can be ungrammatical in French (e.g., *\*est Jean chez lui?*). To address this well-documented learning difficulty, Spada et al. (2005) provided contrastive explicit information about question formation in French and English that aimed to raise learners' awareness of L1-L2 word order differences for question formation and the meaning implications of these word orders. In other words, learners were instructed that declarative word orders with raising intonation are used to ask yes/no questions in French, but not in English. English instead requires interrogative word orders



(i.e., subject-verb inversion). Learners were taught the following ‘rule of thumb’: “To form a yes/no question using *can*, *will*, *is* and *are*, invert the subject and the auxiliary verb” (p. 210). Explicit instruction was followed by communicative practice in English that required use of the target feature. For example, participants played “Find Someone Who...”, which required learners to ask each other yes/no questions like “can you play hockey” and “are you taking piano lessons?”. Teachers provided crosslinguistic metalinguistic corrective feedback during the communicative practice. Results indicated that this instruction improved the accuracy of learners’ question formation in English, measured immediately after instruction using a grammaticality judgement task and oral and written production of question forms. However, because performance was only assessed immediately after instruction, the durability of this type of instruction is unclear.

Although Spada et al. (2005) provided explicit instruction about L1-L2 differences, that study did not examine whether the instruction actually led to language awareness (for discussion of the links between instruction and awareness, see Roehr-Brackin, 2018). Since language awareness was not assessed, it remains possible that the meaning-based practice led to the observed learning gains rather than the provision of EI. Ammar, Lightbown and Spada (2010) therefore extended this line of research with question formation to specifically examine the extent to which awareness of L1-L2 differences was related to performance on a grammaticality judgment test and a scrambled question test. Awareness of L1-L2 differences and similarities was elicited using an oral interview: learners were asked to explain some of their answers in the judgement and scrambled question test. In addition, learners were directly asked whether “there were any differences between French and English with respect to question formation. If they answered that there were differences, they were asked to explain those differences” (Ammar et

al., 2010, p. 134). Results showed more accurate performance among learners who verbalized awareness of L1-L2 differences, adding support to Spada et al.'s (2005) previous findings and suggesting that L1 awareness of L1-L2 differences can play an important role in learning L2 features with L1-L2 differences.

Although this previous research has mostly examined L1-L2 differences in terms of syntax, Kupferberg (1999; Kupferberg & Olshtain, 1996) examined how raising learners' awareness of L1-L2 differences improved Hebrew-speaking learners' use of the Past Perfect in L2 English. This study therefore sought to improve learners' functional knowledge of verbal forms using EI about L1-L2 form-meaning differences. Instruction required learners to translate Hebrew sentences into English, which was followed by metalinguistic contrastive EI about structural and functional L1-L2 differences. Written production results showed that EI about L1-L2 structural and functional differences improved the accuracy of L2 production. Although Kupferberg (1999) did not discuss the precise nature of the EI, the activities at least suggest that attention was paid to the meanings of Hebrew and English verbal morphology to express past perfect information. Furthermore, while short-term benefits were found, no delayed posttesting was included to assess the durability of these instructional effects.

Informed by these lines of previous research about the benefits of explicit instruction about L1-L2 form-meaning differences for improving L2 grammatical learning, McManus and Marsden (2017, 2018, 2019a, 2019b) recently investigated the extent to which these findings might also benefit English-speaking learners' use of the French *Imparfait*, which presents well-documented learning difficulties due to crosslinguistic differences and polyfunctionality (Howard, 2005; Kihlstedt, 2015; McManus, 2013, 2015). As previously discussed, L1-L2 form-meaning mapping differences for habituality and ongoingness are understood to be an important

source of the L2 learning problem (Howard, 2005; McManus, 2015). McManus and Marsden (2017, 2018, 2019b) examined the extent to which EI about L2 and L1 form-meaning mappings for past habituality and past ongoingness improved L2 learning outcomes. Instruction was comprehension-based and included different types of EI and comprehension practice. An ‘L2-only group’ received EI about L2 form-meaning mappings plus comprehension practice of L2 sentences. An ‘L2+L1 group’ received the same L2 EI and L2 comprehension practice plus additional EI about L1 form-meaning mappings and comprehension practice of L1 sentences. An ‘L2+L1prac group’ also received the same as the L2+L1 except no L1 EI was received. Instruction lasted 3.5 hours and was delivered over four weeks (for more information about the study design, see Method and McManus & Marsden, 2017). Results showed that EI about L1 and L2 form-meaning mappings followed by comprehension practice of French (L2) and English (L1) sentences improved learners’ speed (online self-paced reading test) and accuracy (offline sentence judgement task in reading and listening) of aspectual interpretation (*Imparfait, Passé Composé, Présent*) four days after instruction (Immediate Posttest) and six weeks later (Delayed Posttest).

While McManus and Marsden’s (2017, 2018) post-instruction evidence suggested that additional L1 EI and practice benefited L2 performance during untimed judgement tests and real-time sentence processing tests, very little is understood about the extent to which performance on these tests related to learners’ awareness of L1-L2 differences and similarities for this target feature, understood to be an important factor in explaining L2 performance in offline tests (Ammar et al., 2010). In other words, the extent to which this instruction actually led to or improved learners’ awareness of L1-L2 differences remains unknown. The current study investigated this question. No previous research has investigated the extent to which (a) explicit

instruction about L1 and L2 leads to the creation/development of awareness about L1 and L2 form-meaning mappings and (b) the extent to which this awareness relates to language use.

## **Current study**

The current study addresses the following gaps in previous research. First, the extent to which differences in the type of EI (about L2 form-meaning mappings vs. L2 and L1 form-meaning mappings) and comprehension practice (of L2 sentences vs. L2 and L1 sentences) resulted in awareness of L1 and/or L2 form-meaning mapping differences and similarities. Second, how these types of language awareness related to performance in a variety of language tests, thus extending previous research showing that explicit instruction about L1 and L2 form-meaning mappings benefitted L2 performance. The current study sought to address the following research questions:

1. To what extent does providing explicit instruction (EI plus comprehension practice) lead to awareness of form-meaning mappings as measured by verbal reports?
2. To what extent does language awareness resulting from instruction relate to post-instruction L2 performance in comprehension and oral production tests?

## **Method**

### ***Participants***

Participants were 69 university learners of French as a foreign language in semester two of a four-year Bachelor of Arts Honours degree program in England. All participants were aged 18-21, had completed A2-level French<sup>i</sup>, and had not spent more than six weeks abroad in a

French-speaking country. Mean years learning French was 10.3 (SD=2.7) and the mean time spent abroad in a French-speaking country was 3.3 weeks (SD=6.1).

### ***Design***

The study included three testing phases (Pretest in week 1, Posttest in week 5, Delayed Posttest in week 12) and four groups (Test-only, L2-only, L2+L1, L2+L1prac). All treatments were administered via E-Prime 2.0 (Schneider, Eschman & Zuccolotto, 2012) one-to-one using laptops in a lab. Participants were assigned to a group using matched randomization based on Pretest performance, resulting in 16 in the Test-only group, 17 in the L2-only group, and 17 in the L2+L1 group. 19 participants were in the L2+L1prac group. Treatments were delivered to the L2-only, L2+L1, and L2+L1prac groups in four 45-minute sessions over three weeks: two sessions in week one, and one session each in weeks two and three. Each session had a different instructional focus: present vs. past ongoingness (session 1), present vs. past habituality (session 2), past ongoingness vs. past habituality (session 3), and past ongoingness vs. past habituality vs. past perfectivity (session 4). The Test-only group received no treatment and only completed outcome measures at Pretest, Posttest, and Delayed Posttest. Explicit French grammar instruction as part of learners' program of study took place one semester prior to this study only and not during the study, corroborated by interviews with university tutors. The whole study was piloted on a condensed timescale with 10 comparable learners from a different university.

### ***Target feature***

All exemplars of IMP were third-person singular forms: 25 regular (e.g., *jouait* ‘play’) and 23 irregular (e.g., *finissait* ‘finish’) verb types balanced across 48 lexical verb types: twelve states (e.g., be happy), twelve activities (e.g., walk in the park), twelve accomplishments (e.g., run to the shop) and twelve achievements (e.g., find a penny). Verb type frequency was balanced across the four lexical semantic classes using Lonsdale and Le Bras’s (2009) frequency dictionary of French. Only third-person singular forms were selected to maintain the instructional focus on learning IMP’s habitual and ongoing meanings (and not other inflections, for example, which would have expanded the instructional focus to include subject-verb agreement). IMP was additionally balanced for verb regularity, verb lexical type, and verb frequency because these factors can influence IMP learning and use (for review, see Salaberry, 2008). Lack of experimental control on these variables had the potential to threaten the study’s internal validity by introducing confounds (see Mackey & Gass, 2015).

### ***Instructional treatments***

All three instructional treatments included an identical core of L2 EI and comprehension practice (for extensive description of these treatments, see also McManus & Marsden (2017) and IRIS for all materials). This common core is briefly presented before describing the L1 components uniquely received by the L2+L1 and L2+L1prac groups. A description of the EI and comprehension practice used in Session 1 is provided in the supplementary materials.

*L2 EI.* EI about the L2 was pre-practice, provided for approximately five minutes at the start of each session, and during the comprehension practice following incorrect answers (see Appendix for EI used in Session 1). Following cognitive approaches to SLA (e.g., Tyler, 2012a), short videos and images were used to depict conceptual information (e.g., ongoingness),

understood to be more effective for depicting complex concepts than the use of words. For example, in Session 1, the concept of ongoingness was depicted using a short video of a man eating an apple bite by bite, but the apple never gets fully eaten. After seeing the video, participants were asked to think about (but not verbalize) how they might express in French what they just saw in the video. Two possibilities were provided: *il mange une pomme* ('he is eating an apple') and *il mangeait une pomme* ('he was eating an apple'). Recommendations to aid processing were then provided. For example, attend to the verb ending to distinguish present from past ongoingness (-e vs. -ait in writing, mǎʒ vs. mǎʒɛ in speech [the EI used audio recordings for speech, not IPA]).

*L2 comprehension practice.* Pre-practice EI was followed by form-meaning mapping comprehension practice of French sentences, in equal amounts of listening and reading, that required learners to attend to the meanings expressed by IMP, *Présent* and *Passé Composé* to complete the task (i.e., verbal inflections were 'task-essential', see Loschky & Bley-Vroman, 1993). For example, Session 1's aim was for learners to interpret IMP and *Présent* inflections to distinguish present ongoingness from past ongoingness, so learners first read or heard a French sentence (e.g., *il joue au foot* 'he plays/is playing football') and then had to select the stimulus's meaning from two options (e.g., 'right now' vs. 'in the past') (see Table 2 for examples of the L2 and L1 sentences used in the comprehension practice).

<TABLE 2 >

The L2 practice included 552 exemplars (384 in IMP [192 habitual, 192 ongoing], 96 in *Présent*, 72 in *Passé Composé*), balanced across reading and listening. Aural stimuli were

recorded by two L1 French speakers. The French sentences were verified for authenticity and comprehensibility by 26 L1 French speakers.

*L2+L1 treatment.* In addition to the same L2 EI and L2 practice, the L2+L1 treatment included brief EI about English form-meaning mappings for habituality and ongoingness, lasting approximately 3 minutes, which followed the same design as the L2 EI (see Appendix for L1 EI used in session 1). The aim of the L1 EI was to increase learners' awareness to (a) the concepts of ongoingness and habituality and (b) the ways in which English expresses these concepts. For instance, in Session 1 (present vs. past ongoingness), learners saw the same man-eating-the-apple video as used in the L2 EI session and were asked to think about how they might express in English what they just saw in the video. Two possibilities were provided: *he is eating an apple* and *he was eating an apple*. Recommendations to aid processing were then provided. For example, attend to the verb auxiliary (*is* vs. *was*) to distinguish present from past ongoingness.

The L1 comprehension practice followed the same design features as described for the L2 comprehension practice, but with fewer sentences: 160 English sentences (56 in Past Progressive [ongoing], 56 in Past Simple [habitual], 16 in Present Simple [habitual], 16 in Present Progressive [ongoing], 16 in Past Simple [perfective], equally balanced across reading, listening, and lexical aspect type). See Table 2 for examples of the L1 practice.

*L2+L1prac treatment.* This was very similar to the L2+L1 treatment, except that participants received no EI about English, neither before nor during the practice. Participants completed the same L1 comprehension practice as in the L2+L1 treatment.

### ***Outcome tests***



Two versions of each test were administered in a split-block design to reduce test familiarity effects between consecutive test points (e.g., test version A at Pretest and Delayed Posttest, and test version B at Posttest). All outcome tests are available from IRIS.

### *Sentence Completion Test with source of knowledge probes*

The Sentence Completion Test (SCT) provided participants with an image of an event situated in the past followed by a grammatical French sentence with the inflected verb removed and placed in brackets in its infinitive form. The participant was required to read the sentence aloud and inflect the verb so that it appropriately described the image (see Figure 1). This design and procedure was adapted from Lee and Kim (2007), who provided participants with an image and a fill-in-the-blanks text, requiring participants to write in the missing verbal inflections to appropriately describe the image. The current study required learners to orally complete the SCT to reduce potential planning and monitoring effects that a written version might encourage.

The test included sixteen past habitual sentences (requiring IMP), fourteen past ongoing sentences (requiring IMP), and fourteen past perfective sentences (requiring PC). Each sentence had its own image (in contrast to Lee & Kim, 2007). Verbs were balanced across lexical classes, verb regularity, and frequency.

After completing each sentence, participants were then asked “what is the basis of your answer?”, explained at beginning of the test as how did you choose what verb form/tense to use. Participants selected one option out of the following: “a rule”, “something I remember”, “a feeling/intuition”, “a guess” (see Rebuschat, 2013). When “a rule” or “something I remember”

was selected, participants were asked to say the rule and/or what they remembered. The decision to orally elicit responses about rules/feeling/guess was based on previous research that also used oral tests because written tests might discourage fuller responses (see Ammar et al., 2010).

The SCT with source of knowledge probes was administered using PowerPoint, one-to-one with a researcher, and audio-recorded. It took approximately fifteen minutes to complete. Figure 1 in an excerpt from the SCT.

[Figure 1 here]

### *Context-Matching tests in reading and listening*

The Context-Matching tests provided participants with an English context that described either a habitual or an ongoing activity and a French stimulus sentence that either matched or mismatched the meaning of the English context. The French stimulus was always shorter than the English context and, importantly, the French stimuli and English contexts were not translations of each other. The context gave a fuller description of an event in which either a habitual or ongoing function of IMP would be required. As a result, the Context-Matching tests (CMTs) did not elicit direct translations between context and stimulus, but the specific meaning implications of the target feature (IMP). For example:

<b>English Context</b>	Yesterday, Patrick was expecting his wife to come back from
(ongoing):	work any minute. Just as he was on his way out, she appeared in
	the driveway.

**French stimulus**                      *Quand Patrick quittait la maison, il a vu sa femme*

(ongoing):                                “When Patrick was leaving the house, he saw his wife”

Two CMTs were administered: first a listening (CMT-Listen), then a reading (CMT-Read), each with 24 target and 8 filler sentences. In half of test items, the context and stimulus matched (e.g., ongoing context with ongoing stimulus), and in the other half they were mismatched (e.g., habitual context with ongoing stimulus). The CMTs drew on 24 of the 48 lexical verbs (3<sup>rd</sup> person singular) from the intervention<sup>ii</sup>. Items were counterbalanced across the match and mismatch conditions for: ongoingness/habituality, verb frequency, lexical aspect class, verb regularity, and clause ordering (main->subordinate / subordinate-> main).

In both reading and listening CMTs, the English context appeared on screen for 10 seconds. Then, the French stimulus was presented aurally (CMT-Listen) or as text (CMT-Read). Participants were instructed to rate how good the match was between the meaning of the French stimulus and English context by pressing a number on the keyboard from 1 (“very good”), 2 (“good”), 3 (“neither good nor bad”), 4 (“poor”), 5 (“very poor”), with a separate option of 9 for “I don’t know”. The written French stimulus remained on screen until a number was pressed, and then participants could not change their answer. The task was untimed and took approximately 20-25 minutes to complete.

In addition to the pilot study, the English contexts were checked with three native speakers of English, the French stimuli with 26 native speakers of French, and the match and mismatch combinations with three L1 English very advanced learners of French.

*Self-Paced Reading test*

The self-paced reading (SPR) test was administered after the CMTs and used 16 items from the CMT-Listen<sup>iii</sup>, with eight context-stimulus matches and eight mismatches, counterbalanced as described for the CMT tests. Half the items were followed by Yes/No comprehension questions to increase the likelihood that participants focused on meaning (see Keating & Jegerski, 2015)<sup>iv</sup>. The answers to the questions depended on a lexico-semantic feature (but not verb stems) and not on inflectional morphology. For each trial, the English context appeared for 10 seconds before an X appeared in the centre of the screen. A spacebar press brought up the first and then each subsequent word of the French stimuli. After the last word, the next screen displayed “END”. Participants were instructed to read as quickly as possible. Reaction times were collected from each word according to the noncumulative moving-window procedure (Marinis, Roberts & Felser, 2005). The font was 18-point Courier New, displayed in the centre of a white background, without line breaks.

### **Data coding and analysis**

For the SCT, rules were coded following Izquierdo and Collins (2008). Four types of rule were coded. First, a rule was coded as ‘adverbial’ when adverbs were used as a cue for verbal morphology choices (e.g., “I used the Imperfect because it goes with *hier* (‘yesterday’) [participant 201a]. Second, a rule was coded as ‘form-function mapping’ when explicit reference was made to the meaning implications of French tense forms, but without reference to L1 (e.g., “The Imperfect is used with habits and things that are repeated” [participant 301b]. Third, a rule was coded as ‘L1-L2 form-function mapping’ when explicit reference was made to L1-L2 differences at the level of form-function mappings (e.g., “I used *Imparfait* because it’s ongoing and that’s different to English” [participant 267b]. Fourth, a rule was coded as ‘L1 translation’

when the learner referred to direct translation from the L1 (e.g., “I translated ‘he walked’ in my head first and then did it like that” [participant 232c]).

For the CMTs, responses were coded as follows: correct responses were scored five points (e.g., selecting “very good” or “good” for matched test items), incorrect responses were scored 1 point (e.g., selecting “poor” or “very poor” for matched test items), and midway responses were scored three points (e.g., selecting “neither good nor bad” for matched test items).

For the SPR tests, a “critical word” in each French stimulus disambiguated ongoing from habitual meaning of the IMP and thus determined whether the sentence matched or mismatched the English context. The critical word in ongoing sentences was the auxiliary of the *Passé Composé* (e.g., *Il quittait la maison quand il a vu sa femme* ‘he was leaving the house when he saw his wife’) and in habitual sentences it was IMP (e.g., *Il quittait la maison quand il voyait sa femme*, ‘he would be leaving the house when he saw his wife’). For more information, see McManus and Marsden (2017, 2018). RTs for the critical word were calculated from the onset of the critical word to the onset of the next word. Raw RT data were analyzed, removing critical word RTs less than 150 ms and greater than 2000 ms (Keating & Jegerski, 2015).

Descriptive statistics and graphics were examined for normality of distribution, which indicated neither normal distribution nor equal variances (according to box plots, Q-Q plots, and Shapiro-Wilks tests). As a result, 4 x 3 x 3 robust repeated measures (RM) ANOVAs with bootstrapped procedures were carried out. The between-subjects factors were Group (L2+L1, L2+L1prac, L2-only, Test-only) and knowledge source (Rule, Feeling, Guess) and test point was the within-subjects factor (Pretest, Posttest, Delayed Posttest). The alpha level was set at .05 (but

interpretation of results is based on effects sizes and 95 % confidence intervals of the effect size, not the p-value, see below).

To examine relationships between language awareness and L2 learning, individuals were coded as ‘L2 aware’, ‘L1-L2 aware’, or ‘unaware’ based on their rule responses in the SCT at Delayed Posttest (as described above). If more than 30% of rules included reference to L1 and L2 form-meaning mappings, that learner was coded as ‘L1-L2 aware’. If 30% of rules referenced L2 form-meaning mappings without reference to L1, however, that learner was coded as ‘L2 aware’. If no rules or fewer than 30% of rules mentioned form-meaning mappings, that learner was coded as ‘unaware’. Although previous research has tended to code awareness based on a single mention of the feature under study (for review, see Leow, 2015), an awareness threshold of 30% of mentions or greater was used to provide a more stringent classification. This is because a single mention could refer to item-specific knowledge that does to characterize a more general rule and its application (see Roehr-Brackin, 2018). Requiring a higher cut-off indicates that those coded as ‘L1-L2 aware’ or ‘L2 aware’ repeatedly reported ‘L1-L2’ or ‘L2’ knowledge in their completion of the test.

In terms of performance on the outcome tests, gains scores between Pretest and Delayed Posttest were computed (i.e. Delayed Posttest score minus Pretest score). In order to examine relationships between types of language awareness (L1-L2 aware, L2 aware, unaware) and grammatical learning, biserial correlations with bootstrapped confidence intervals were calculated (Field, 2013).

Cohen’s *d* and Hedges’ *g* effect sizes (ES) and 95% confidence intervals (CIs) for *d* and *g* were used to interpret magnitudes of change for all between- and within-subjects paired comparisons. Within-subject ES were calculated using the mean and standard deviation of the

Pretest as a baseline (and the Posttest for effect sizes at Delayed Posttest). CIs that did not include zero were considered reliable indicators of change (Field, 2013). Effect size values were interpreted using Plonsky and Oswald's (2014) field-specific benchmarks: Cohen's *d* within-subjects: 0.60 (small), 1.00 (medium), 1.40 (large); Cohen's *d* between-subjects: .40 (small), .70 (medium), 1.00 (large); *R* values: .25 (small), .40 (medium), and .60 (large).

## **Results**

### ***Source of knowledge in the Sentence Completion test***

#### *Rules, intuitions, and guesses*

Table 3 shows each group's source of knowledge responses on the SCT at Pretest, Posttest, and Delayed Posttest. Although 'rule' is the most frequent response type for all of the groups, 'feeling' emerges as an alternative source of knowledge response for the L2+L1 group at Posttest and Delayed Posttest. No other group selects 'feeling' as a source of knowledge response.

<TABLE 3>

There was a statistically significant three-way interaction between Group x Time x Knowledge Source ( $F_{5,4,117.4} = 10.2, p = .000, \eta^2 p = .320$ ), indicating that not all groups performed similarly in their selection of knowledge source types over time. To further examine these effects, between-group posthoc comparisons of source of knowledge responses were

conducted at Pretest, Posttest, and Delayed Posttest (see Table 4). At Pretest, there were no between-group differences (negligible ES). At Posttest, however, the L2+L1 group provided fewer rule-based responses compared to all other groups (large ES) and more feeling-based responses (large ES). There were no differences between L2+L1prac and L2-only on any knowledge source type (negligible ES). Furthermore, although Test-only showed only negligible differences compared to L2+L1prac, the L2-only group provided more rule-based responses (medium ES) and fewer guesses (medium ES) than Test-only. This patterning of results was repeated six weeks later at Delayed Posttest: negligible differences between (a) L2+L1prac and L2-only and (b) L2+L1prac and Test-only. The L2-only group provided more rule-based responses (large ES) and fewer guesses (large ES) than Test-only, but this time to a large extent. Compared to all other groups, L2+L1 group provided fewer rule-based responses (large ES) and more feeling-based responses (large ES).

In terms of within-group performance over time (see Table 5), there was no change over time (Pre-Post, Pre-Delayed, Post-Delayed) for Rule, Feeling, or Guess for both Test-only and L2-only (negligible ES), meaning that selection of the different rule types did not change meaningfully over time for learners in these groups. The L2+L1prac group made more rule-based selection at Posttest than at Pretest (small ES) and made fewer guesses at Posttest than at Pretest (small ES), but no Pretest-Delayed or Post-Delayed differences were found. When compared with Pretest performance, the L2+L1 group made fewer rule-based selections at Posttest (small ES) and at Delayed Posttest (large ES) and more feeling-based selections at Posttest and Delayed Posttest (large ES for both). Guesses also reduced between Pretest and Posttest (medium ES). No reliable changes were found between Posttest and Delayed Posttest for any knowledge source type (negligible ES).



Taken together, these results indicate that, for the L2+L1 group, the proportion of feeling/intuition responses increased and rules/guess decreased as explanations for their performance in the SCT. Trends similar to these were not found for the other groups.

<Table 4>

<Table 5>

### *Types of rules*

Since rules constituted a frequent source of knowledge choice, learners' explanations for selecting a rule were also analysed (see Table 6). Participants reported that four types of rule were used when deciding what verbal form to use: use of adverbs, L2 form-function mappings, L2 and L1 form-function mappings, and L1 translation. There was a statistically significant three-way interaction between Group x Time x Rule Explanation ( $F_{12.2, 255} = 13.03$ ,  $p = .000$ ,  $\eta^2p = .383$ ), indicating that not all groups selected the same types of rules.

<Table 6>

To further examine these effects, between-group posthoc comparisons of source of knowledge responses were conducted at Pretest, Posttest, and Delayed Posttest (see Table 7). Pretest results showed no between-group differences (negligible ES). At Posttest, results showed that adverb-based explanations were significantly higher for Test-only than for all other groups (large ES). This is because form-function-based rules instead of adverb-based rules increased

among the treatment groups. However, two clear trends are evident: (a) reference to L2 form-function mappings and (b) reference to L2 and L1 form-function mappings. Compared to L2+L1 and Test-only, L2 form-function mappings were frequent among L2+L1prac and L2-only learners (medium-to-large ES), with no differences between these groups (negligible ES). In contrast, only the L2+L1 group repeatedly referenced L2-L1 form-function mappings, which represents an important difference between L2+L1 and all other groups (large ES). L1 translation rules were absent among L2+L1 learners (medium-to-large ES), but were used among L2+L1prac, L2-only, and Test-only, with no differences between these groups (negligible ES). A very similar patterning of results was found at Delayed Posttest, indicating little change between Posttest and Delayed Posttest. Although adverb use was infrequent among the treatment groups, it was used more frequently in the Test-only group (large ES). Only the L2+L1 group reported using L2-L1 form-function explanations.

In terms of within-group comparisons (see Table 8), there was no change over time for Test-only in the use of any rule type (negligible ES). An important change for all treatment groups was a large reduction in adverb-based rules between Pretest-Posttest and Pretest-Delayed Posttest (large ES). L2 form-function rules increased for all treatment groups between Pretest-Posttest and Pretest-Delayed Posttest (large ES). Exceptionally for the L2+L1 group, use of L1 and L2 form-function mapping rules increased between Pretest-Posttest and Pretest-Delayed Posttest (large ES). Furthermore, while L1 translation explanations reduced between Pre-Post and Pre-Delayed among L2+L1 learners (large ES), no change over time was found among L2+L1prac and L2-only learners (negligible ES).

<Table 7>

<Table 8>

### ***Language awareness and L2 learning gains***

Learners were re-grouped according to whether their rule-based explanations at Delayed Posttest included reference to form-function mappings. As previously noted, a learner was coded as ‘aware’ when  $\geq 30\%$  of rule explanations referred to form-meaning mappings (L2 and/or L2+L1), excluding L1 translation (see Table 9). This awareness coding resulted in the following groupings: 26 participants demonstrated awareness of L2 form-meaning mappings, 16 participants demonstrated of L2+L1 form-meaning mappings, and 27 participants demonstrated no awareness of form-meaning mappings.

<Table 9>

These new ‘awareness’ groupings were then used to examine relationships between language awareness, specifically whether learners were aware of the meaning implications of verbal morphology in L2 only or in L2 and L1, and performance on four different outcome measures tapping different types of language performance abilities: self-paced reading, CMT-read, CMT-listen (from McManus & Marsden, 2017, 2018) and accuracy in the SCT. Pretest-Delayed Posttest gains scores for these measures are shown in Table 10.

Descriptively, these results show that compared to L2 aware learners, L2+L1 aware learners demonstrated faster processing (SPR,  $g$  [CIs] = -1.42 [-2.13, -.74]), more accurate interpretation in listening (CMT-Listen,  $g$  [CIs] = 1.71 [.98, 2.43]) and reading (CMT-Read,  $g$

[CIs] = .93 [.27, 1.58]), and more accurate oral production (SCT,  $g$  [CIs] = 1.11 [.44, 1.77]) of French morphosyntax.

To further explore relationships between awareness and learning gains, biserial correlations were conducted between awareness of L2+L1 form-meaning mappings and Pre-Delayed performance gains. This analysis showed positive correlations between L2+L1 awareness and performance on each measure that indicated positive, medium-to-large sized relationships between L2+L1 awareness and language performance: SPR ( $r_b = .588$ ), CMT-Read ( $r_b = .415$ ), CMT-Listen ( $r_b = .636$ ), SCT ( $r_b = .491$ ). See also supplementary materials for  $p$ -values and CIs for  $r_b$  statistics and scatter plots. Furthermore, using the coefficient of variation ( $R^2$ ), these results additionally indicate that L2+L1 awareness accounts for between 40% and 17% of the variability in learning gains: 34.6% (SPR), 17.2% (CMT-Read), 40.4% (CMT-Listen), 24.1% (SCT).

<Table 10>

## **Discussion**

The present study examined the extent to which EI about L2 and L1 form-meaning mappings plus comprehension practice of L1 and L2 sentences resulted in language awareness of L1 and L2 form-meaning mappings. Language awareness was then related to performance in four outcome tasks: a self-paced reading test, two untimed comprehension tests (listening, reading), and an oral production test.

### ***Summary of results***

Pretest results showed that use/selection of a verbal form (e.g., IMP) was initially guided by adverbs, but this rule reduced over time for all treatment groups (but not the Test-only group). In other words, learners reported less reliance on adverb presence/absence as an important cue for verbal morphology selection following treatments that focused on the meaning implications of verbal forms. In addition, all treatments resulted in greater awareness of form-meaning relationships. However, only learners receiving the L2+L1 treatment (i.e. the only treatment that included EI about L1 form-meaning mappings) reported awareness of L1-L2 form-meaning mapping differences/similarities. Finally, in terms of correlations between language awareness and Pretest-Delayed Posttest learning gains, quicker processing and more accurate performance in reading, listening, and speech were found only for learners who indicated awareness of L1-L2 form-meaning mappings.

### ***A role for explicit information about L1 form-meaning mappings***

Providing EI about L1 form-meaning mappings (i.e. the L2+L1 treatment) appeared to raise awareness about L1 processing routines, including L1-entrenched attention and competing cues (Ellis 2006; MacWhinney, 2012). This awareness additionally appeared to benefit L2 performance. Learners from no other group demonstrated language improvement in these ways. It is likely that this additional EI about the L1 was more helpful than EI about L2 due to the nature of the crosslinguistic learning problem: L1-L2 form-meaning mappings differences for habituality and ongoingness. It is possible that the aspectual function of *-ed* was too difficult to induce from practice alone. One explanation for this difficulty could be the low saliency and the multiple and complex form-meaning mappings for *-ed* because *-ed* can express both perfective and habitual meaning. L1 EI may have helped learners to clarify and/or establish more efficiently

the meanings expressed by *-ed* and how these meanings could, in turn, be expressed in French. In short, the EI may have better helped learners (a) work out the meanings expressed by English verbal forms and (b) how these same meanings are expressed in L2.

Indeed, Pretest performance in the SCT indicates a lack of general awareness about the meaning implications expressed by French verbal forms. When asked about why a particular form was used, learners frequently reported adverb cues. For many learners, therefore, knowing when to use a particular aspectual form (e.g., IMP) appeared to be influenced by adverbial information rather than the meaning implications of that particular verbal form. Interviews with teachers additionally confirmed that adverb co-indexation was a common explanation, also known to be common in textbooks. For example, one teacher noted that “[...] some words and phrases are nearly always used with the *Imparfait*, while others seem to stick like glue to the *Passé Composé*. I use adverb lists to help students determine which tense to use [...]” (Teacher 5).

An important finding from the current study, therefore, is that instruction emphasising the meaning implications of verbal forms can be helpful for L2 learning. It also appears that instruction focusing on how the same meanings were expressed in L1 and L2 was particularly helpful. The finding that adverb-based cues were common, especially at Pretest, indicates that learners were not aware of the meaning implications of aspectual forms in French, even though they were relatively experienced classroom learners (mean years learning French was 10.3). This finding that meaning-based instruction was helpful for learning aspectual forms in L2 French, however, contrasts with some previous research (e.g., Harley, 1989; Izquierdo, 2014; Mifka-Profozic, 2015). Two explanations could account for this difference with previous research. First, informed by work in cognitive linguistics (e.g., Tyler, 2012a), the current study provided

meaning-based instruction that used videos and images to depict meanings (e.g., ongoingness). It is possible that text-based descriptions and discussions of the meaning implications of aspectual forms as used in previous research are less transparent and more difficult to conceptualize than the use of videos and images. Second, the current study reversed the typical grammar instruction model that introduces grammatical forms and then discusses the meanings they express (e.g., L'Huillier, 1999). In the current study, a meaning was introduced first (e.g., ongoingness) followed by discussion of the grammatical forms that could express this meaning (e.g., IMP). It is possible that increasing learners' awareness to meaning first was more helpful than introducing forms first, especially for these relatively experienced classroom learners who were familiar with verbal forms but not their meanings (as evidenced by Pretest performance). The extent to which the current study's learning gains can be attributed to these specific instructional differences is an empirical question, however. Very little research has systematically examined differences in the nature of explanations and explicit information in L2 grammatical learning.

Lastly, while the current study's findings appear to contrast with some previous research in L2 aspect instruction, they are consistent, however, with research demonstrating the importance of instruction that increases learners' awareness of the meaning implications of words and sentences (e.g., Tyler, 2012a; Tyler & Ortega, 2018). While this state-of-affairs necessarily challenges approaches to grammar instruction that privilege form over meaning, it also challenges growing interest in 'comprehensible input' approaches to language teaching that emphasise almost exclusive L2 use in the classroom without developing learners' awareness and understanding of language (see ACTFL, 2014). This is despite recommendations from national language teaching associations that developing learners' awareness and understanding of

language, including of L1-L2 differences, is necessary for language development (see NSFLEP, 2015).

### ***Language awareness and L2 learning gains***

The finding that awareness of L1-L2 form-meaning mappings correlated with improved performance in a variety of tests is consistent with previous research, particularly with Ammar et al. (2010) for accuracy and McManus and Marsden (2017, 2018) for offline and online comprehension. Indeed, these findings are also consistent with research showing that metalinguistic knowledge, in general, can be beneficial for improved L2 performance (Roehr-Brackin, 2018). While a growing body of research indicates that awareness of the target feature can be helpful for improving learning outcomes amongst instructed L2 learners, the current study suggests that the nature of that awareness appears important (likely determined by the nature of the learning problem). In other words, awareness about L1 properties may be most useful for target features exhibiting L1-L2 differences. This, however, is an empirical question for future research.

Although McManus and Marsden (2017, 2018) did not correlate language awareness measurements with performance, those findings showed that performance following instruction about L2 and L1 led to more accurate and faster processing than for learners receiving L2 EI only. Furthermore, these benefits were also maintained at Delayed Posttest, six weeks later. Taken together, these results indicate that (a) EI about L1 and L2 led to increased awareness of L1 and L2 form-meaning mappings for the target feature and (b) that this awareness contributed to learners' improved performance in a variety of different tests. These findings therefore indicate an important role for awareness about L1 and L2 form-meaning mappings in improving



language performance understood to be influenced by L1-L2 differences at the level of form-meaning mappings.

### ***Limitations and future research***

Given the small sample sizes, these accounts are tentative and require replication (Porte & McManus, 2019). In addition, the grouping of learners based on reported awareness ('L2 aware', 'L2-L1 aware', 'unaware') was imbalanced because these groupings were based on test performance following instruction. While this approach is more favourable than assuming that all learners in the L2-only group were aware of L2 form-meaning mappings only, for example (see Leow, 2015, for discussion), the between-group imbalance is a limitation. This limitation was addressed in the interpretation of any findings by using Hedges' *g* effect sizes (appropriate for imbalanced group comparisons, see Plonsky, 2015) and the confidence intervals for *g*. Despite these limitations, understanding L1 awareness in L2 grammatical learning constitutes an important research agenda if we wish to understand the cognitive processes underpinning learning and the ways in which instruction might be tailored to optimize potential learning outcomes.

In terms of future research directions, this study's findings and implications (as well as those reported in McManus & Marsden, 2017, 2018, 2019a, 2019b) need to be extended to other languages, target structures, and learner populations to more fully understand the role of EI about L1 and L1 awareness in L2 learning. It seems that the effectiveness of EI about L1 will likely depend on the nature of the target feature. It is also an empirical question whether morphosyntactic features like grammatical gender (that arguably lack meaning) would also benefit from this type of instruction. As such, future research is needed to help verify the current

study's findings and the extent to which they are generalizable to other target features. For example, research indicates that English speakers struggle with assigning subject-object information in L2 German (Jackson, 2007) and L2 Spanish (Fernandez, 2008) given that word order is a reliable cue for assigning subject-object roles in English, but not in German and Spanish. Examining the extent to which EI about L1 for subject-object information can benefit L2 learning is one way to extend the current findings to new target features with the same learning problem: L1 and L2 differ in how they express the same meanings.

In addition, a potentially helpful future research direction informed by this study's interviews with teachers would be to better understand L1 use and L1 awareness techniques in world language classrooms. As previously noted, teachers indicated that some of the learning difficulties they face in the classroom could be related to learners' lack of awareness of their own language (see also McMillan & Turnbull, 2009). Although some attempts are being made to incorporate L1 and L2 use into world language classrooms (see NSFLEP, 2015), we understand very little about teachers' views of L1 and L2 use in the classroom, especially for less commonly taught languages. Furthermore, we understand very little about the frequencies and functions of L1 and L2 use in these classrooms (but see Polio & Duff, 1994). Future research should systematically examine both teachers' attitudes and beliefs towards L1 use in world language classrooms as well as document (e.g., creation of oral corpora) language use in a variety of world classrooms, including of less commonly taught languages. These avenues would help better understand relationships between language awareness and L2 learning and teaching.

## References

- Amenós-Pons, J., Ahern, A., & Fuentes, P. G. (2017). L1 French learning of L2 Spanish past tenses: L1 transfer versus aspect and interface issues. *Studies in Second Language Learning and Teaching*, 7, 489-515.
- American Council on the Teaching of Foreign Languages (2014). *The Language Educator*, 9(5), 1-68.
- Ammar, A., Lightbown, P. M., & Spada, N. (2010). Awareness of L1/L2 differences: does it matter?. *Language Awareness*, 19(2), 129-146.
- Andersen, R. W. (1984). The one to one principle of interlanguage construction. *Language Learning*, 34, 77–95.
- Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the Passé Composé and Imparfait. *The Modern Language Journal*, 85(2), 226-243.
- Ayoun, D. (2013). *The second language acquisition of French tense, aspect, mood and modality*. Amsterdam: John Benjamins
- Bartning, I. (2009). The advanced learner variety: Ten years later. In E. Labeau & F. Myles (Eds.), *The Advanced Learner Varieties: The Case of French* (pp. 11–40). Bern, Switzerland: Peter Lang.
- Cintrón-Valentín, M. C., & Ellis, N. C. (2016). Salience in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology*, 7, 1284.
- Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge, UK: Cambridge University Press.
- DeKeyser, R. M. (1997). Beyond explicit rule learning: Automating second language morphosyntax. *Studies in Second Language Acquisition*, 19(2), 195-221.

- Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27(2), 164-194.
- Ellis, N. C., & Sagarra, N. (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition*, 32(4), 553-580.
- Fernández, C. (2008). Reexamining the role of explicit information in processing instruction. *Studies in Second Language Acquisition*, 30, 277-305.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. London, UK: Sage.
- Hoffmann, L.-F. (1995). *L'essentiel de la grammaire française*. Englewood Cliffs, NJ: Prentice Hall
- Horst, M., White, J., & Bell, P. (2010). First and second language knowledge in the language classroom. *International Journal of Bilingualism*, 14(3), 331-349.
- Howard, M. (2005). Les contextes prototypiques et marqués de l'emploi de l'imparfait par l'apprenant du français langue étrangère. In E. Labeau & P. Larrivée (Eds.), *Nouveaux développements de l'imparfait* (pp. 175–197). New York/Amsterdam: Rodopi.
- Izquierdo, J. (2014). Multimedia instruction in foreign language classrooms: Effects on the acquisition of the French perfective and imperfective distinction. *Canadian Modern Language Review*, 70(2), 188-219.
- Izquierdo, J., & Collins, L. (2008). The facilitative role of L1 influence in tense–aspect marking: A comparison of Hispanophone and Anglophone learners of French. *The Modern Language Journal*, 92(3), 350-368.

- Jackson, C. N. (2007). The use and non-use of semantic information, word order, and case markings during comprehension by L2 learners of German. *The Modern Language Journal*, 91(3), 418-432.
- Kihlstedt, M. (1998). *La référence au passé dans le dialogue: Etude de l'acquisition de la temporalité chez des apprenants dits avancés de français*. University of Stockholm: Akademitryk.
- Kihlstedt, M. (2015). Acquisition of the imparfait in L2 French in adults and children: The same or different?: A comparative case study of Swedish adults and children in an immersion setting. *Language, Interaction and Acquisition*, 6(1), 74-106.
- Keating, G.D., & Jegerski, J. (2015). Experimental designs in sentence processing research: A methodological review and user's guide. *Studies in Second Language Acquisition*, 37(1), 1-32.
- Kupferberg, I. (1999). The cognitive turn of contrastive analysis: Empirical evidence. *Language Awareness*, 8, 210-222.
- Kupferborg, I., & Olshtain, E. (1996). Explicit contrastive instruction facilitates the acquisition of difficult L2 forms. *Language Awareness*, 5, 149-165.
- Lee, E., & Kim, H. Y. (2007). On crosslinguistic variations in imperfective aspect: the case of L2 Korean. *Language Learning*, 57(4), 651-685.
- Leow, R. P. (2015). *Explicit Learning in the L2 Classroom: A Student-Centered Approach*. New York: Routledge.
- L'Huillier, M. (1999). *Advanced French Grammar*. Cambridge: Cambridge University Press.
- Lonsdale, D., & Le Bras, Y. (2009). *A frequency dictionary of French: Core vocabulary for learners*. Routledge.

- Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In Crookes, G. & Gass, S.M. (eds). *Tasks and Language Learning. Integrating Theory and Practice* (pp. 123-167). Bristol, UK: Multilingual Matters
- Mackey, A., & Gass, S.M. (2015). *Second language research: Methodology and design*. New York: Routledge.
- MacWhinney, B. (2012). The logic of the unified model. In Gass, S.M. & Mackey, A. (eds). *The Routledge Handbook of Second Language Acquisition* (pp. 211-227). New York: Routledge
- Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27(1), 53-78.
- McManus, K. (2015). L1-L2 differences in the acquisition of form-meaning pairings: A comparison of English and German learners of French. *Canadian Modern Language Review*, 71(2), 51-77.
- McManus, K. (2013). Prototypical influence in second language acquisition: what now for the Aspect Hypothesis? *International Review of Applied Linguistics in Language Teaching*, 51(3), 299-322.
- McManus, K. & Marsden E. (2019a). Signatures of automaticity during practice. Explicit instruction about L1 processing routines can improve L2 grammatical processing. *Applied Psycholinguistics*, 40(1), 205-234.
- McManus, K. & Marsden, E. (2019b). Using explicit instruction about L1 to reduce crosslinguistic effects in L2 grammar learning. Evidence from oral production in L2 French. *The Modern Language Journal*, 103(2)

- McManus, K. & Marsden, E. (2018). Online and offline effects of L1 practice in L2 grammar learning. A partial replication. *Studies in Second Language Acquisition*, 40(2), 459-475.
- McManus, K. & Marsden, E. (2017). L1 explicit instruction can improve L2 online and offline performance. *Studies in Second Language Acquisition*, 39(3), 459-492
- McMillan, B. & Turnbull, M. (2009). Teachers' use of the first language in French immersion: Revisiting a core principle. In Turnbull, M. & Dailey-O'Cain, J. (eds). *First Language Use in Second and Foreign Language Learning* (pp. 15-34). Bristol, UK: Multilingual Matters.
- Mitchell, R.F., Tracy-Ventura, N., & McManus, K. (2017). *Anglophone Students Abroad: Identity, Social Relationships and Language Learning*. New York: Routledge
- Mifka-Profozic, N. (2015). Effects of corrective feedback on L2 acquisition of tense-aspect verbal morphology. *Language, Interaction and Acquisition*, 6(1), 149-180.
- National Standards in Foreign Language Education Project (NSFLEP). (2015). *World-Readiness Standards For Learning Languages*. Alexandria, VA: ACTFL
- Plonsky, L. (2015). Statistical power, p values, descriptive statistics, and effect sizes: A “back-to-basics” approach to advancing quantitative methods in L2 research. In Plonsky, L. (ed). *Advancing Quantitative Methods in Second Language Research* (pp. 43-65). New York: Routledge.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878-912.
- Polio, C. G., & Duff, P. A. (1994). Teachers' language use in university foreign language classrooms: A qualitative analysis of English and target language alternation. *The Modern Language Journal*, 78(3), 313-326.

- Porte, G. & McManus, K. (2019). *Doing Replication Research in Applied Linguistics*. New York: Routledge
- Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63(3), 595-626.
- Rebuschat, P. (Ed.). (2015). *Implicit and Explicit Learning of Languages*. Amsterdam, NL: John Benjamins.
- Roehr-Brackin, K. (2018). *Metalinguistic Awareness and Second Language Acquisition*. New York: Routledge.
- Salaberry, R. M. (2008). *Marking past tense in second language acquisition: A theoretical model*. New York, NY: Continuum.
- Slobin, D. (1973). Cognitive prerequisites for the development of grammar. In *Ferguson, C.A. & Slobin, D.I. Studies of Child Language Development* (pp. 175-208). New York, NY: Holt, Rinehart & Winston
- Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime 2.0: User's guide*. Pittsburgh: Psychology Software.
- Smith, C. (1997). *The parameter of aspect*. 2nd. Netherlands: Kluwer Academic Publishers.
- Spada, N., & Lightbown, P. M. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *The Modern Language Journal*, 83(1), 1-22.
- Spada, N., Lightbown, P. M., & White, J. (2005). The importance of form/meaning mappings in explicit form-focused instruction. In Housen, A. & Pierrard, M. (eds). *Investigations in Instructed Second Language Acquisition* (pp. 199-234). New York: Gruyter.
- Tagliamonte, S., & Lawrence, H. (2000). "I Used to Dance, but I Don't Dance Now" The Habitual Past in English. *Journal of English Linguistics*, 28, 324-353.



- Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning*, 64(2), 279-309.
- Tyler, A. (2012a). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. New York: Routledge.
- Tyler, A. (2012b) Spatial Language, Polysemy, and Cross-Linguistic Semantic Mismatches: Cognitive Linguistics Insights into Challenges for Second Language Learners. *Spatial Cognition & Computation: An Interdisciplinary Journal*, 12(4), 305-335.
- Tyler, A. & Ortega, L. (2018). Usage-inspired L2 instruction: An emergent, researched pedagogy. In Tyler, A., Ortega, L., Uno, M., & Park, H. I. (Eds.). *Usage-inspired L2 Instruction: Researched Pedagogy*. Amsterdam, NL: John Benjamins.
- VanPatten, B. (2015). Foundations of processing instruction. *International Review of Applied Linguistics in Language Teaching*, 53(2), 91-109.
- Zhao, H., & MacWhinney, B. (2018). The Instructed Learning of Form–Function Mappings in the English Article System. *The Modern Language Journal*, 102(1), 99-119.

TABLES

TABLE 1.

Past habituality and past ongoingness in French and English

Meaning	French sentence with English gloss
Past habituality	<i>Elle mangeait une baguette (e.g., tous les jours)</i> ‘She ate / would eat / used to eat a baguette (everyday)’
Past ongoingness	<i>Elle mangeait une baguette (e.g., quand le téléphone a sonné)</i> ‘She was eating a baguette (when the telephone rang)’

TABLE 2

Examples of L2 and L1 Sentences Used in the Comprehension Practice from Session 1 (English glosses provided for illustration)

Target meaning	Present ongoing	Past ongoing
French stimulus used in L2 practice (received by all treatment groups)	<i>Il...</i> ‘He’	<i>Il...</i> ‘He’
	<i>joue au tennis</i> ‘is playing tennis’	<i>jouait au tennis</i> ‘was playing tennis’
	<i>porte une cravate</i> ‘is wearing a tie’	<i>portait une cravate</i> ‘was wearing a tie’
English stimulus used in L1 practice (received by L2+L1 and L2+L1prac groups)	<i>He...</i>	<i>He...</i>
	<i>is eating a sandwich</i>	<i>was eating a sandwich</i>
	<i>is knocking at the door</i>	<i>was knocking at the door</i>
Response options	<i>Maintenant [X]</i> ‘Now’ <i>Dans le passé</i> ‘In the past’	<i>Maintenant</i> ‘Now’ <i>Dans le passé [X]</i> ‘In the past’

TABLE 3.  
 Source of knowledge responses on Sentence Completion test (mean percentages and SDs)

	L2+L1	L2+L1prac	L2-only	Test-only
	M(SD)	M(SD)	M(SD)	M(SD)
Pretest				
Rule	91.7 (9.4)	93 (6.7)	95.5 (9.4)	93.9 (8.6)
Feeling	0 (0)	0 (0)	0 (0)	0 (0)
Guess	8.3 (9.4)	7 (6.7)	4.5 (9.4)	6.1 (8.6)
Posttest				
Rule	84.5 (7.1)	97.6 (3.3)	98.5 (3.2)	94.9 (4.9)
Feeling	15.5 (7.1)	0 (0)	0 (0)	0 (0)
Guess	0 (0)	2.4 (3.3)	1.5 (3.2)	5.1 (4.9)
Delayed				
Rule	79 (7.6)	96.6 (3.9)	99.6 (1.07)	95.7 (4.2)
Feeling	21 (7.6)	0 (0)	0 (0)	0 (0)
Guess	0 (0)	3.4 (3.9)	0.4 (1.07)	4.3 (4.2)

TABLE 4

Between-group comparisons for knowledge source at Pretest, Posttest and Delayed Posttest (p, Cohen's *d* ES)

	Pretest			Posttest			Delayed Posttest		
	Rule <i>p, d</i>	Feeling <i>p, d</i>	Guess <i>p, d</i>	Rule <i>p, d</i>	Feeling <i>p, d</i>	Guess <i>p, d</i>	Rule <i>p, d</i>	Feeling <i>p, d</i>	Guess <i>p, d</i>
L2+L1 v L2+L1prac	>.05, -.16	>.05, 0	>.05, .16	.00, -2.41	.00, 3.18	>.05, -1.21	.00, -2.97	.00, 4.03	.01, -1.20
L2+L1 v L2-only	>.05, -.40	>.05, 0	>.05, .40	.00, -2.54	.00, 3.09	>.05, -.66	.00, -3.41	.00, 3.91	>.05, -.53
L2+L1 v Test-only	>.05, -.24	>.05, 0	>.05, .24	.00, -1.70	.00, 3.04	.00, -1.50	.00, -2.70	.00, 3.85	.00, -1.47
L2+L1prac v L2-only	>.05, -.31	>.05, 0	>.05, .31	>.05, -.28	>.05, 0	>.05, .28	>.05, -1.02	>.05, 0	.02, 1.02
L2+L1prac v Test-only	>.05, -.12	>.05, 0	>.05, .12	>.05, .66	>.05, 0	>.05, -.66	>.05, .22	>.05, 0	>.05, -.22
L2-only v Test-only	>.05, .18	>.05, 0	>.05, -.18	>.05, .88	>.05, 0	.02, -.88	>.05, 1.29	>.05, 0	.00, -1.29

*Note.* Grey shading = Cohen's *d* effect size equal to or greater than .40 with CIs for *d* that do not pass through zero (indicative of a meaningful difference). See supplementary materials for CIs.

TABLE 5.

Within-group comparisons for knowledge source at Pretest, Posttest and Delayed Posttest (p, Cohen's d ES)

	Pretest-Posttest			Pretest-Delayed Posttest			Posttest-Delayed Posttest		
	Rule <i>p, d</i>	Feeling <i>p, d</i>	Guess <i>p, d</i>	Rule <i>p, d</i>	Feeling <i>p, d</i>	Guess <i>p, d</i>	Rule <i>p, d</i>	Feeling <i>p, d</i>	Guess <i>p, d</i>
L2+L1	.03, .86	.00, -3.09	.00, 1.25	.00, 1.49	.00, -3.91	.00, 1.25	>.05, .75	>.05, -.75	>.05, 0
L2+L1prac	.01, -.87	>.05, 0	.01, .87	>.05, -.66	>.05, 0	>.05, .66	>.05, .28	>.05, 0	>.05, -.28
L2-only	>.05, -.43	>.05, 0	>.05, .43	>.05, -.61	>.05, 0	>.05, .61	>.05, -.46	>.05, 0	>.05, .46
Test-only	>.05, -.14	>.05, 0	>.05, .14	>.05, -.27	>.05, 0	>.05, .27	>.05, -.18	>.05, 0	>.05, .18

*Note.* Grey shading = Cohen's *d* effect size equal to or greater than .60 with CIs for *d* that do not pass through zero (indicative of a meaningful change). See supplementary materials for CIs.

TABLE 6.

Rule types used in sentence completion test (mean percentages and SDs)

	L2+L1	L2+L1prac	L2-only	Test-only
	M(SD)	M(SD)	M(SD)	M(SD)
<b>Pretest</b>				
Adverb	62.4 (20.7)	61.4 (29.1)	59.1 (23.9)	60.6 (17.8)
L2 Form-Function	17.1 (23.7)	20.4 (11.2)	16 (8.7)	18.8 (14.3)
L2+L1 Form-Function	0 (0)	0 (0)	0 (0)	0 (0)
L1 translation	20.5 (11.2)	18.2 (26.4)	24.4 (21.8)	19.4 (6.3)
Other	0 (0)	0 (0)	0.5 (1.9)	1.2 (3.3)
<b>Posttest</b>				
Adverb	8.6 (7.1)	6.8 (10.3)	5.8 (5.6)	65 (14.4)
L2 Form-Function	55.6 (22.1)	77.9 (31.8)	74.6 (13.9)	17.8 (7.5)
L2+L1 Form-Function	35.6 (24.6)	0	0	0
L1 translation	0	15.3 (25.1)	19.4 (10.7)	16.9 (11.4)
Other	0.2 (0.8)	0 (0)	0.2 (0.7)	0.3 (1.2)
<b>Delayed</b>				
Adverb	2.8 (4.8)	8.5 (9.5)	7.6 (12.7)	61.1 (19.2)
L2 Form-Function	57.5 (16.9)	69.8 (22.9)	77.8 (22.6)	17.9 (14)
L2+L1 Form-Function	39.7 (15.5)	0 (0)	0	0
L1 translation	0 (0)	19.3 (15.1)	14.4 (13.1)	19.6 (15.1)
Other	0 (0)	2.4 (10.4)	0.2 (0.8)	1.4 (3.9)

Table 7.

Between-group comparisons for types of rules used at Pretest, Posttest and Delayed Posttest (p, Cohen's d ES)

	Pretest				Posttest				Delayed Posttest			
	Adverb	L2 Form-Function	L2-L1 Form-Function	L1 translation	Adverb	L2 Form-Function	L2-L1 Form-Function	L1 translation	Adverb	L2 Form-Function	L2-L1 Form-Function	L1 translation
	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>	<i>p, d</i>
L2+L1 v L2+L1prac	>.05, .04	>.05, -.18	>.05, 0	>.05, .11	>.05, .20	.02, -.81	.00, 2.11	.03, -.84	>.05, -.74	>.05, -.61	.00, 3.73	.00, -1.76
L2+L1 v L2-only	>.05, .15	>.05, .06	>.05, 0	>.05, -.23	>.05, .44	>.05, -1.03	.00, 2.05	.00, -2.56	>.05, -.50	.02, -1.02	.00, 3.62	.01, -1.55
L2+L1 v Test-only	>.05, .09	>.05, -.09	>.05, 0	>.05, .12	.00, -5.02	.00, 2.26	.00, 2.01	.02, -2.13	.00, -4.23	.00, 2.54	.00, 3.57	.00, -1.87
L2+L1prac v L2-only	>.05, .09	>.05, .44	>.05, 0	>.05, -.25	>.05, .12	>.05, .13	>.05, 0	>.05, -.21	>.05, .08	>.05, -.35	>.05, 0	>.05, .35
L2+L1prac v Test-only	>.05, .03	>.05, .13	>.05, 0	>.05, -.06	.00, -4.72	.00, 2.50	>.05, 0	>.05, -.08	.00, -3.57	.00, 2.68	>.05, 0	>.05, -.02
L2-only v Test-only	>.05, -.07	>.05, -.24	>.05, 0	>.05, .31	.00, -5.48	.00, 5.04	>.05, 0	>.05, .23	.00, -3.31	.00, 3.16	>.05, 0	>.05, -.37

*Note.* Grey shading = Cohen's *d* effect size equal to or greater than .40 with CIs for *d* that do not pass through zero (indicative of a meaningful difference). See supplementary materials for CIs.

Table 8.  
 Within-group comparisons for types of rules used at Pretest, Posttest and Delayed Posttest ( $p$ , Cohen's  $d$  ES)

	Pretest-Posttest				Pretest-Delayed Posttest				Posttest- Delayed Posttest			
	Adverb	L2 Form- Function	L2-L1 Form- Function	L1 translation	Adverb	L2 Form- Function	L2-L1 Form- Function	L1 translation	Adverb	L2 Form- Function	L2-L1 Form- Function	L1 translation
	$p, d$	$p, d$	$p, d$	$p, d$	$p, d$	$p, d$	$p, d$	$p, d$	$p, d$	$p, d$	$p, d$	$p, d$
L2+L1	.00, 3.48	.00, -1.68	.00, -2.05	.00, 2.59	.00, 3.97	.00, -1.96	.00, -3.62	.00, 2.59	.04, .96	>.05, -.10	.00, .20	>.05, 0
L2+L1prac	.00, 2.50	.00, -2.41	>.05, 0	.>.05, .11	.00, 2.44	.00, -2.74	>.05, 0	>.05, -.05	>.05, -.17	>.05, .29	>.05, 0	>.05, -.19
L2-only	.00, 3.07	.00, -5.05	>.05, 0	>.05, .29	.00, 2.69	.00, -3.61	>.05, 0	>.05, .56	>.05, -.18	>.05, -.17	>.05, 0	>.05, .42
Test-only	.>.05, -.27	>.05, .09	>.05, 0	>.05, .27	>.05, -.03	>.05, .06	>.05, 0	>.05, -.02	>.05, .23	>.05, -.01	>.05, 0	>.05, -.20

*Note.* Grey shading = Cohen's  $d$  effect size equal to or greater than .60 with CIs for  $d$  that do not pass through zero (indicative of a meaningful difference). See supplementary materials for CIs.



TABLE 9.

Participants indicating form-meaning awareness in their explanations for rules (raw frequencies and group percentages)

	L2 form-meaning awareness				L2+L1 form-meaning awareness		
	Pre	Post	Delayed		Pre	Post	Delayed
L2+L1 (n=17)	0	15 (88%)	16 (94%)		0	15 (88%)	16 (94%)
L2+L1prac (n=19)	0	16 (84%)	14 (74%)		0	0	0
L2-only (n=17)	0	14 (82%)	12 (70%)		0	0	0
Test-only (n=16)	0	0	0		0	0	0

*Note.* Participant coded as aware when  $\geq 30\%$  of rule explanations refer to form-meaning mappings (L2 and/or L2+L1), excluding L1 translation

Table 10.

Gains scores (means, SDs) between Pretest and Delayed Posttest in self-paced reading, listening and reading judgement, and sentence completion tests

	N	SPR gains (ms)	CMT-reading gains (accuracy, max = 5)	CMT-listening gains (accuracy, max = 5)	SCT completion gains (% accuracy)
L2 awareness	26	-52.5 (131.6)	0.3 (1.0)	0.3 (0.7)	26.9 (14.8)
L2+L1 awareness	16	-222.6 (85.1)	1.1 (0.5)	1.4 (0.5)	47.4 (22.7)
No awareness	27	54.6 (330.4)	0.3 (0.8)	-0.1 (0.7)	6.6 (24.9)

*Note.* SPR, CMT-read, and

---

<sup>i</sup> An advanced-level, English school leaving qualification that assesses learners' French language abilities in speaking, reading, writing, listening, and grammar, as well as cultural and social knowledge of French-speaking countries. It is equivalent to CEFR level B2, typically after 700-800 hours of instruction. See <https://www.aqa.org.uk/subjects/languages/as-and-a-level/french-7652/specification-at-a-glance>

<sup>ii</sup> The same lexical verbs were used in each version, but in different contexts (e.g., he went to the shop vs. she went to the stadium)

<sup>iii</sup> So that participants had not read the same stimuli twice in the same test phase.

<sup>iv</sup> Judgements about match/mismatches were not requested because explicit judgements in online tasks can lead to unusually slow or careful reading as well as tapping into more explicit resources (see Keating & Jegerski, 2015).