

**USING EXPLICIT INSTRUCTION ABOUT L1 TO REDUCE CROSSLINGUISTIC
EFFECTS IN L2 GRAMMAR LEARNING: EVIDENCE FROM ORAL PRODUCTION
IN L2 FRENCH**

Kevin McManus, Pennsylvania State University, USA, kmcmanus@psu.edu

Emma Marsden, University of York, UK, emma.marsden@york.ac.uk

ABSTRACT

This study advances previous research about the effects of explicit instruction on second language (L2) development by examining learners' use of verbal morphology following different types of explicit information (EI) and comprehension practice. We investigated the extent to which additional EI about L1 can reduce the effects of crosslinguistic influence in L2 oral production. Sixty-nine English-speaking learners of L2 French undertook either: (a) a 'core' treatment of EI about the L2 with L2 comprehension practice, (b) the same L2 core + L1 comprehension practice, (c) the same L2 core + L1 comprehension practice + EI about L1, or (d) outcome tests only. Results showed that providing additional EI about the L1 benefitted the accuracy of oral production immediately after the instruction and then 6 weeks later. These results suggest that tailoring instruction, specifically the nature of the EI, to the nature of the learning problem can facilitate L2 learning. In particular, EI about L1 can facilitate L2 learning by increasing learners' awareness of similarities and differences in how L1 and L2 express the same meanings.

Keywords: crosslinguistic influence; foreign language learning; French; instruction; oral production; first language; grammar

A major focus of second language acquisition (SLA) research to date has sought to understand the competition and relationships between a learner's different languages (Calabria et al., 2018). This research has repeatedly shown that use of a single language activates a speaker's other known languages (Marian & Spivey, 2003; Wu & Thierry, 2010), that prior first language (L1) knowledge and experience can influence second language (L2) use (e.g., selective attention to linguistic cues, Ellis & Sagarra, 2011; MacWhinney, 2012), and that L1-L2 differences can influence the route and rate of L2 morphosyntactic development and processing (Avery & Marsden, under review; Isabelli, 2008; McManus 2013, 2015; Murakami, 2016; Roberts & Liszka, 2013). However, despite major advances in what we know about the cognitive effects and mechanisms of learning a second language, little research has systematically examined the next step in this program: how can this understanding about the competition and relationships between a learner's different languages be used to facilitate L2 learning and teaching?

Although explicit instruction remains a dominant approach in classrooms for reducing crosslinguistic influence during L2 learning (Ranta & Lyster, 2017), the extent to which it can actually benefit L2 morphosyntactic development constitutes a long-standing debate (for reviews, see DeKeyser, 2017; VanPatten, 2017). One line of research contributing to this debate has compared practice with and without explicit information (EI) about the L2 target feature. While some research has shown that practice with EI provides few if any learning benefits compared to practice in making connections between forms and their meanings (Marsden 2006; Sanz & Morgan-Short, 2004; VanPatten & Oikkenon, 1996), others have found that EI about the L2 appears to play an important role by drawing learners' attention to specific aspects of the target feature, thus enhancing the effectiveness of the practice (Fernández, 2008; Henry, Jackson, & DiMidio, 2017; VanPatten et al., 2013).

These lines of investigation have helped us to understand the effectiveness of EI about the L2. However, a notable consistency in this previous research is the adoption of presence/absence designs, examining, for example, broad effects of practice with and without EI (e.g., Andringa, de Glopper, & Hacquebord, 2011; Sanz & Morgan-Short, 2004; Tolentino & Tokowicz, 2014) or with and without comprehension (or production) practice (DeKeyser & Sokalski, 1996; Li & DeKeyser, 2017). Less research has manipulated the nature of EI given in the instruction to address specific crosslinguistic learnability challenges. One exception is McManus and Marsden (2017, 2018, 2019), who manipulated the type of EI (and comprehension) practice across conditions. The current study addresses this gap by introducing L1 EI and L1 practice into L2 instruction.

We compared three types of EI and comprehension practice designed to improve English-speakers' use of the *Imparfait* (IMP) in L2 French, a target feature well-documented to be late-acquired due to complex L1-L2 form-meaning mapping differences (Bartning, 1997, 2009; Howard, 2005; Kihlstedt, 2015; McManus 2013, 2015): one group received EI about the L2 plus extensive comprehension practice of L2 sentences; a second group received the same L2 EI, L2 comprehension practice, plus additional comprehension practice of L1 sentences; and a third group received the same L2 EI, L2 comprehension practice, L1 comprehension practice, plus additional L1 about the EI. This design allowed us to compare (a) EI about L2 form-meaning mappings with EI about both L2 and L1 form-meaning mappings and (b) comprehension practice of L2 sentences with comprehension practice of both L2 and L1 sentences. Of particular interest was the extent to which explicit instruction about the L1 can address learning difficulties resulting from L1-L2 form-meaning mapping differences.

We begin by reviewing SLA research about crosslinguistic influence in L2 grammar learning and follow this with an overview of research that has investigated EI about the L2 to

address crosslinguistic influence. The extent to which EI about the L1 may be able to improve L2 learning is then briefly reviewed.

BACKGROUND

Crosslinguistic Influence in L2 Grammatical Learning

Research to date has repeatedly shown that a speaker's prior linguistic knowledge and experience can influence L2 grammatical learning in two specific ways. First, the same L1 and L2 linguistic cues (e.g., verbal inflections, word order) can vary in the meanings they index (MacWhinney, 2012). Second, prior language knowledge and experience can influence attention to cues (Ellis, 2006, 2008; Wulff & Ellis 2018).

The Unified Competition Model (MacWhinney, 2012) proposes that crosslinguistic influence can be at least partly determined by the 'availability' and 'reliability' of linguistic cues in L1 and L2. Linguistic cues can vary in type (morphological, syntactic, prosodic, semantic, and pragmatic), availability (how frequently cues are present), reliability (how often cues lead to the same interpretation), and validity (the joint product of availability and reliability). The Unified Competition Model predicts crosslinguistic influence when the validity of the same cue differs crosslinguistically. One linguistic feature exemplifying this learning problem because of crosslinguistic variation is viewpoint aspect (Smith, 1997), a semantic category that expresses how speakers present or view events in time (Comrie 1976; Dahl & Velupillai, 2013): Past perfective viewpoints present events as complete (e.g., she ran to the park [yesterday]), past habitual viewpoints present events as regularly repeated (e.g., she ran to the park [everyday], she used to run to the park [everyday]), and past ongoing viewpoints present events as in progress (e.g., she was running to the park [yesterday]). All languages can express these viewpoint aspect

meanings (Smith 1997), but they differ in how they map them to forms. One language pairing that maps viewpoint meanings differently is English and French (see Table 1 for examples):

1. English uses Simple Past to express both past perfectivity and past habituality (*Simple Past*, Comrie, 1976; Tagliamonte & Lawrence, 2000), but French expresses these meanings by using a different verbal form for each meaning (*Passé Composé* for perfectivity; IMP for past habituality (Hoffmann, 1995).
2. French uses IMP to express both past habituality and past ongoingness (Hoffmann, 1995), but English uses a different verbal inflection to express each of these meanings (*Simple Past* for habituality; *Past Progressive* for ongoing, Comrie, 1976; Tagliamonte & Lawrence, 2000).

Cue validities for viewpoint aspect in English and French are therefore different because of inconsistent mappings between viewpoint aspect meanings and linguistic cues. These differences are hypothesized to give rise to crosslinguistic influence (MacWhinney 2012).

Furthermore, compared to past habituality in English, which can be indexed by a variety of linguistic cues (predominately Simple Past, but also *would*, *used to*, and temporal adverbials like ‘everyday’ Tagliamonte & Lawrence, 2000), past ongoingness in English is indexed by one linguistic cue (Past Progressive). This variation of past-habituality-indexing cues lowers the validity of English cues for habituality relative to those for ongoingness (see also Andersen, 1984; Slobin, 1973). It is hypothesized that a meaning indexed by low validity cues in the L1 reduces sensitivity to that meaning, which, as a result, delays learning of L2 cues indexing that meaning. In other words, the English speaker is predicted to be less sensitive to the concept of habituality because the L1 cues indexing this meaning are multiple and of low reliability (see also Athanasopoulos & Bylund, 2013; McManus, 2015; Slobin, 1973). Therefore, greater

learning difficulties are predicted for learning IMP’s habitual function compared to its ongoing function because of the low validity of L1 cues for habituality.

<TABLE 1 HERE>

TABLE 1
Viewpoint Aspect Meanings in French Sentences with English Glosses

Viewpoint meaning	French sentence with English gloss
Past habituality	<i>Elle jouait au foot (e.g., tous les jours)</i> ‘She played / would play / used to play football (everyday)’
Past ongoingness	<i>Elle jouait au foot quand le telephone a sonné</i> ‘She was playing football when the telephone rang’
Past perfectivity	<i>Elle a joué au foot (hier)</i> ‘She played football (yesterday)’

Crosslinguistic Influence in L2 Learning of IMP

SLA research shows patterns of learning associated with IMP’s different viewpoint aspect meanings that can be attributed to different cue validities in L1 and L2 (Ayoun, 2004, 2013; Howard, 2005; Kihlstedt, 1998; McManus 2013, 2015). Given that IMP is used to express both past ongoingness and past habituality, research indicates (a) that these viewpoint aspect meanings are not acquired together and (b) that the acquisition order of these meanings appears to be influenced by L1 background: Ongoingness acquired before habituality for English-speaking learners (Howard, 2005), but habituality acquired before ongoingness for Swedish-speaking learners (Kihlstedt, 1998). These observations suggest that the configuration of L1 form-meaning mappings could play a role in explaining IMP acquisition (see also Andersen, 1984; Ayoun, 2013, MacWhinney, 2012; Salaberry, 2008).

Focusing on habituality, English speakers have been shown to initially use the past perfective *Passé Composé* to express habituality (e.g., “*parfois je suis allée* visiter mes amis à Paris le weekend*” [sometimes I went^{*-PAST PERFECTIVE} to visit my friends in Paris at the weekend], instead of *je visitais* [I visited^{-PAST HABITUAL}]; from Howard, 2005, p. 188) while using IMP appropriately to express ongoingness (Howard, 2005; McManus 2015, see also Ayoun, 2004; Starren 2001). This usage reflects (a) how English maps viewpoint aspect meanings to verbal forms (i.e., L2 learners express habituality using a past perfective form because their L1 does this) and (b) a need to grammatically distinguish one meaning from the other, as done in their L1: one form for ongoingness (*Past Progressive*) and a different form for habituality (*Simple Past*).

An important question informing these lines of research is the extent to which instruction tailored to the nature of the learning problem (e.g., increasing learners’ sensitivity to the concept of habituality and the L1 and L2 cues that index it) can facilitate learning in cases of persistent crosslinguistic influence brought about both by low cue validity in the L1 and different cue validities between L1 and L2. In the following section, we review research designed to reduce persistent crosslinguistic influence effects in L2 grammatical learning. We focus on two main approaches: (a) explicit instruction about L2, (b) explicit instruction about L2 and L1.

Explicit Instruction about the L2

An important body of work informed by theoretical and empirical research about persistent crosslinguistic influence effects in L2 learning, especially for polyfunctional forms such as French IMP, has examined the extent to which instruction that addresses the cause of crosslinguistic influence can improve L2 learning (e.g., Cintrón-Valentín & Ellis, 2016; Ellis & Sagarra, 2011; VanPatten, 2017). One approach to this has provided EI about language

processing strategies (i.e. information about cues, what cues to attend to) followed by practice in order to develop more appropriate L2 processing behaviours (e.g., Henry, Jackson, & DiMidio, 2017; Tolentino & Tokowicz, 2014; Zhao & MacWhinney, 2018).

Based on evidence that extensive prior use of the L1 tunes how speakers attend to language and subsequently biases which cues get noticed and processed (Ellis, 2006, 2008; Wulff & Ellis 2018), Ellis and colleagues used comprehension practice with (correct/incorrect) feedback to manipulate attention to cues that might be missed due to L1-L2 cue validity differences (Cintrón-Valentín & Ellis, 2016; Ellis & Sagarra, 2011; Ellis et al., 2014). Results at immediate posttest indicated that using explicit instruction to increase attention to L2 cues that would have been missed due to entrenched L1 processing behaviours can improve L2 grammatical learning. Ellis and Sagarra's (2011) meta-analysis of this body of research additionally indicated a graded effect explained by L1-L2 cue validity differences: Chinese speakers (no tense morphology) were found to be less able than speakers of Spanish and Russian (rich tense morphology) to learn L2 inflectional cues in an inflectionally rich language (Latin). This body of research indicates that explicit instruction about L2 targeting (a) competing cues and (b) learned attention resulting from prior language use can improve L2 grammatical processing.

Explicit Instruction about the L1

In a recent review of language pedagogy research, R. Ellis and Shintani (2014) note that “there is almost no research that has investigated the actual effects of the classroom use of the L1 on L2 learning” (p. 247). Albeit a very small body of research, some studies have investigated EI about the L1 to address learning difficulties arising from crosslinguistic influence, specifically for lexis (Laufer & Girsai, 2008; White & Horst, 2012) and grammar (Horst, White & Bell,

2010; Kupferberg, 1999; Spada, Lightbown & White, 2005). These studies have compared interventions consisting of explicit, contrastive information about L1 and L2 with interventions of explicit information about L2 only. For example, to improve French-speaking learners' use of possessive determiners in L2 English, Spada et al. (2005) provided EI about L1 and L2 highlighting that in French a possessive determiner agrees with the grammatical gender of the noun, but in English it agrees with the natural gender of the possessor. Learners were provided with 'rule of thumb' EI: "Ask "Whose is it?" If it belongs to a man or a boy, use his. If it belongs to a woman or girl, use her". This EI was followed by classroom-based, communicative oral practice. For example, learners "played a game in which they had to describe their classmates without using their names: his hair is short and his t-shirt is yellow [...]" (p. 211). Immediate posttest results showed increased accuracy of possessive determiner use in writing and speaking and better verbalization of rules about when and how to use English possessive determiners. Similar benefits were reported by Kupferberg (1999) for Hebrew-speaking learners' use of viewpoint aspect cues in L2 English. 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 learners' production of grammatical aspect forms, especially past perfect (for similar results, see also Kupferberg & Olshtain, 1996).

These lines of research indicate benefits for providing EI about L1 and L2 combined with output practice. However, this research agenda still has some gaps. First, since no delayed posttests were used and tests (largely) elicited language of a more controlled nature (rather than under time and oral communicative pressure), the durability and generalizability of learning gains remains unclear. Second, these studies have not addressed more complex cases of

crosslinguistic influence at the level of form-meaning mappings arising from L1-L2 cue validity differences, like IMP use among English-speaking learners of French.

We additionally observe that this research has only investigated the benefits of L2 practice. For example, although Spada et al. (2005) provided EI about L1 and L2, the practice was in L2 only. Thus, we do not yet know the extent to which practice involving L1 *and* L2 sentences can reduce crosslinguistic influence in L2 learning. For example, following the tenets of Skill Acquisition Theory (DeKeyser, 2017), EI about L1 followed by practice in interpreting the L1 may help develop and consolidate declarative knowledge about the L1 (e.g., the concept of past habituality and its expression), and make L1 processing explicit in a way that serves more accurate L2 processing.

To address these gaps, McManus and Marsden (2017, 2018, 2019) provided EI about L2 and L1 form-meaning mappings for viewpoint aspect in French (L2) and English (L1) and comprehension practice of both L2 and L1 sentences (unlike any of the aforementioned studies) to investigate their effects on immediate and delayed L2 online and offline processing of aspect in L2 French. McManus and Marsden's instruction lasted 3.5 hours and was delivered over four weeks. Results showed that EI about L1 and L2 processing routines followed by comprehension practice of French (L2) and English (L1) sentences improved learners' speed (online self-paced reading test) and accuracy (offline sentence judgement test in reading and listening) of aspectual interpretation (IMP, *Passé Composé*, *Présent*) during the comprehension practice itself and after it, both four days after (Immediate Posttest) and six weeks after (Delayed Posttest).

Although McManus and Marsden's evidence suggested that L1 EI and comprehension practice (combined with L2 EI and L2 comprehension practice) benefited L2 online and offline comprehension, we do not yet know the extent to which it benefitted other skills, such as oral production. In line with calls to better understand the type of language knowledge and skills

resulting from instruction, examining performance in oral production tests following comprehension practice would allow us to understand the extent to which practice can develop different types of language use (e.g., can comprehension practice only benefit performance on comprehension tests). Evaluating instructional effectiveness on tests that are different to the instruction itself and using more than one test is frequently recommended (e.g., Larsen-Freeman, 2015; Lightbown, 2008; Norris & Ortega, 2000), and it can be useful for both pedagogical (e.g., can teaching help language use in a range of contexts) and theoretical reasons (e.g., inform understanding about transfer appropriate processing, implicit/explicit knowledge accounts, skill specificity, roles for input and input processing; see, for example, Marsden, 2006, for French inflectional verb morphology; Marsden & Chen, 2011, for English tense verb morphology; Kasprovicz & Marsden, 2018, for German inflectional case marking; and Shintani & Ellis, 2013, for a review). Thus, the present study set out to examine the extent to which instruction under a particular condition (i.e. comprehension) benefitted language use in a different condition (i.e. oral production).

THE PRESENT STUDY

We examined whether providing L2 learners with different types of EI (about only L2 form-meaning mappings vs. about both L2 + L1 form-meaning mappings) plus comprehension practice (interpreting only L2 sentences vs. interpreting both L2 + L1 sentences) benefited the accuracy of IMP use in oral production outcome measures immediately after instruction and six weeks later, and whether the type of EI and/or comprehension practice moderated performance. This extended our previous research showing that comprehension practice benefitted online and offline performance in comprehension tests. In the present study, we sought to address the following research questions:

- RQ 1. To what extent can providing comprehension-based instruction (EI plus comprehension practice) improve the accuracy of IMP use in L2 oral production immediately after instruction (Posttest) and six weeks later (Delayed Posttest)?
- RQ 2. Compared to L2-only EI plus practice, to what extent are accuracy changes over time different when providing additional L1 practice with and without L1 EI?

METHODOLOGY

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 French at a British university. All participants were L1 (British) English speakers, aged 18-21, had completed A2-level French (English school leaving qualification, equivalent to CEFR level B2, typically after 700-800 hours of instruction). In terms of amounts/types of previous French language learning, participants reported that previous instruction was predominantly classroom-based (mean = 10.3 years, SD=2.7) with very little time spent abroad in a French-speaking country (mean = 3.3 weeks, SD=6.1). No participant reported extra-curricular use of French.

Target Feature

The target feature was French IMP inflectional verbal morphology, a past tense form used to express past habituality and past ongoingness (e.g., *il jouait au foot* - 'he was playing/used to play football'), selected because SLA research has repeatedly shown this form to

be late-acquired due to functional complexity (Ayoun, 2004, 2013), including complex L1-L2 form-meaning differences for viewpoint aspect (Howard, 2005; McManus 2013, 2015; Kihlstedt, 2015). As a reminder (see previous discussion and Table 1), past habituality and past ongoingness are expressed by the same verbal form in French (IMP, Hoffmann, 1995), but by different forms in English (for past habituality, predominately *Simple Past*, but also *would*, and *used to*; for past ongoingness, *Past Progressive*; Comrie, 1976; Tagliamonte & Lawrence, 2000). This English-French form-meaning mapping difference is understood to be a major obstacle affecting IMP learnability (Howard, 2005; McManus, 2013, 2015), not found for learners of L1s that map viewpoint aspect in similar ways (e.g., Spanish-French learners, see Amenós-Pons, Ahren & Guijarro-Fuentes, 2017; Izquierdo & Collins, 2008; Lorenzo, 2002).

In the present study, all exemplars of IMP were third-person singular: 25 regular (e.g., *marcher* ‘walk’) and 23 irregular (e.g., *courir* ‘run’) verb types balanced across 48 lexical verb types: twelve states (e.g., be happy), twelve activities (e.g., run in the park), twelve accomplishments (e.g., walk to the shop) and twelve achievements (e.g., find a letter). Verb type frequency was balanced across these four lexical semantic classes using Lonsdale and Le Bras’s (2009) frequency dictionary of French.

Study Design

The study included three testing points (Pretest in week 1, Posttest in week 5, Delayed Posttest in week 12) and four groups (L2+L1, L2+L1prac, L2-only, Control). All treatments were administered via laptops using E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2012). Participants were assigned to a group using matched randomization based on Pretest performance, resulting in 16 in the Control 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 in four sessions

over three weeks, each lasting approximately 45 minutes: two sessions in week one, and one session each in weeks two and three¹. Each session had a different instructional focus: present vs. past ongoingness (present tense vs. IMP in session 1), present vs. past habituality (present tense vs. IMP in session 2), past ongoingness vs. past habituality (IMP + IMP vs. IMP + *Passé Composé* in session 3), and past ongoingness vs. past habituality vs. past perfectivity (session 4). The Control group only completed the Pretest, Posttest, and Delayed Posttests and received no treatment. Participants received no explicit French grammar instruction as part of their university program during the study, corroborated by interviews with university tutors. The whole study was piloted on a condensed timescale with ten comparable learners.

Instructional Treatments

All three instructional treatments included an identical core of L2 EI and L2 practice (see Appendix for example). This common core is briefly presented before describing the L1 treatment components uniquely received by the L2+L1 and L2+L1prac groups. For materials for all treatments, see IRIS (www.iris-database.org) and McManus and Marsden (2017) for a fuller description.

L2 EI. EI about the L2 was pre-practice, provided for approximately five minutes at the start of each session, and during-practice following incorrect answers (see Appendix for pre-practice EI used in Session 1). The pre-practice EI depicted conceptual information via a short video and images. 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 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 practice sentences).

<TABLE 2 HERE>

TABLE 2

Examples of L2 and L1 Practice Used in Session 1 (English Glosses Included for Illustration)

Target meaning	Present ongoing	Past ongoing
French stimulus used in L2 practice (received by all treatment groups)	<i>Elle...</i> 'She'	<i>Elle...</i> 'She'
	<i>joue au foot</i> 'is playing football'	<i>jouait au foot</i> 'was playing football'
	<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 drinking a glass of wine</i>	<i>was drinking a glass of wine</i>
	<i>is knocking at the door</i>	<i>was knocking at the door</i>
Response options	<i>Maintenant</i> [X] 'Now'	<i>Maintenant</i> 'Now'
	<i>Dans le passé</i> 'In the past'	<i>Dans le passé</i> [X] 'In the past'

The L2 practice included 552 exemplars (384 in IMP [192 ongoing, 192 habitual], 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 viewpoint aspect, 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' sensitivity to (a) the concepts of ongoingness and habituality and (b) the linguistic cues used in L1 to index these concepts. For instance, in Session 1 (present vs. past ongoingness), learners saw the same man-eating-the-apple

video 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 practice followed the same design features as described for the L2 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 exact same L1 practice as in the L2+L1 treatment.

Oral Production Outcome Measures

To examine the extent to which instruction under a particular condition (i.e. comprehension) benefitted language use in a different condition (i.e. oral production), two different oral production tests were used to assess performance following comprehension practice. See IRIS (www.iris-database.org) for the full tests.

Picture-Based Oral Narrative (to elicit habitual IMP). Two picture-based narrative stories, the cat story and the sister story, as used in previous French L2 research (McManus 2015; Mitchell, Tracy-Ventura & McManus 2017) and adapted from Dominguez et al. (2013), were used. Both stories were structurally similar and set in the past, involving unambiguous perfective contexts (for *Passé Composé* use) and habitual contexts (for IMP use). The stories contrasted the protagonists' long-standing daily routines (i.e., past habitual events) with a one-

time event (perfective). For the cat story, pictures show the daily routines of a girl and her pet cat (habitual events), followed by a specific day when the cat escaped (perfective events). For the sister story, two adult sisters talk about recurrent childhood events (habitual events), followed by the events from a specific day of their holiday in Spain (perfective events). Short instructions in English were provided for completing the stories, a series of French lexical prompts to structure the stories, and a list of five French vocabulary items (two nouns and five verbs) for use when retelling the story. Participants were given two minutes to look through the pictures before telling the story. Both stories were piloted for equivalency with ten L2 learners and ten French L1 speakers.

Activity Description Oral Production Test (to elicit ongoing IMP). This test was designed to elicit descriptions of ongoing/interrupted events in the past. Learners were first shown an event in progress (e.g., a car driving down a road), and then shown the same event but with an interruption (e.g., a police officer stopping the car). The learner was asked to say in French what was happening before the intervening event happened (a context for IMP), as shown in Figure 1.

Short instructions were provided at the start of the test. Participants did not see the images before beginning. Two versions were created, each with 28 stimuli, 16 of which depicted ongoing events, equally balanced across the four lexical aspect classes. The remaining twelve events were distractors. Both versions were piloted for equivalency with ten L2 French learners and ten L1 French speakers.

The 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).

<FIGURE 1 HERE>

FIGURE 1

Example Test Item From the Activity Description Oral Production Test



Data Coding and Analysis

All data were digitally recorded and then orthographically transcribed by an expert user of French using CHAT from CHILDES (MacWhinney, 2000) and protocols designed and tested for French SLA (see Marsden, Myles, Rule & Mitchell, 2003 and www.flloc.soton.ac.uk). All transcripts were double checked for accuracy by the first author and one other expert user of French. CHAT transcripts were first automatically tagged for part-of-speech information using the French MOR program, followed by automatic and manual disambiguation of initial part-of-speech taggings using the French POST program (Parisse & Le Normand, 2000). The %VCX program (Dominguez et al., 2013) was used to automatically identify all verbal inflections, which were then manually tagged for aspectual information (IMP, *Passé Composé*, *Présent*, Other), appropriateness of use (Appropriate, Inappropriate), and context (Habitual, Ongoing, Perfective). This tagging enabled automatic analysis of aspectual information. The CLAN command COMBO was used to automatically compute frequency counts for all combinations of form, (in)appropriateness of use, and context. The first author and a research assistant each manually tagged the same 113 transcripts from each outcome test (25% of the total data) using %VCX, compared their codings, and discussed any differences. The first author coded the

remaining files. Cohen's kappa inter-rater reliability coefficients from these codings were .82 for the Picture-Based Narrative and .88 for the Activity Description Test.²

Our analysis of IMP production used the 'target-like use' (TLU) metric (Pica, 1983; Ellis, 1994), which analyses a morpheme's distribution in both appropriate and inappropriate contexts (rather than just in appropriate contexts, as with 'suppliance in obligatory contexts'). TLU was computed using the frequency counts automatically generated by CLAN, as follows: N of appropriate uses / (Total N of appropriate contexts + N of uses in inappropriate contexts). Following Howard (2005) and Kihlstedt (1998), the stative verbs *avoir* (have) and *être* (be) were excluded from our analyses because they are well-documented to be overused and rote-learned. Appropriacy of IMP use was defined as production of IMP to describe habitual and ongoing events, determined according to the obligatory contexts provided by the tests, as previously described. For instance, the use of *Présent* to describe a past habitual event was coded as inappropriate, whereas use of IMP to describe the same event was coded as appropriate. Inaccurate or invented verb endings were discussed by both raters and were coded as invented forms (i.e., inappropriate, scoring zero) unless both raters agreed that they could be structurally identified as IMP, *Passé Composé*, or *Présent*. For example, *a couré* ('ran'), an invented form similar to the target *a couru*, was coded as an appropriate *Passé Composé* because it was structurally similar to the regular *Passé Composé* (present auxiliary + past participle).

Examination of descriptive statistics and graphics showed that the data were neither normally distributed nor had equal variances (according to box plots, Q-Q plots, and Shapiro-Wilks tests). We therefore present the results of 4 x 3 robust repeated measures (RM) ANOVAs with bootstrapped procedures (Larson-Hall, 2014), with Group as the between-subjects factor (L2+L1, L2+L1prac, L2-only, Control) and test point as the within-subjects factor (Pretest, Posttest, Delayed Posttest). We set the alpha level at .05. Although Mauchly's Test of Sphericity

was not statistically significant ($p > .05$), the residual SSCP matrix showed deviations from Sphericity, so a Greenhouse-Geisser correction factor was used. No important deviations from normality and homogeneity of variances for the residuals were discovered. If, according to a robust RM-ANOVA, a statistically significant effect was found, pairwise comparisons with Bonferroni correction were used for the posthoc tests using the Games-Howell test for separate covariance matrices. Eta squared (η^2) and partial eta squared (η_p^2) are reported for all omnibus tests (Norouzian & Plosnky, 2017).

Cohen's d effect sizes (ES) and 95% confidence intervals (CIs) for d were used to interpret magnitudes of change for all between- and within-subjects paired comparisons (instead of p -values, Larson-Hall & Plonsky, 2015). Within-subject ES at Posttest were calculated using the mean and standard deviation of the Pretest as a baseline, and at Delayed Posttest using the Posttest as baseline. CIs for d that included zero were considered unreliable indicators of change (Field, 2013). We also calculated between-group ES changes with effects adjusted for baseline differences, that provide similar information to 'gains scores', and present these in supplementary materials (see McManus & Marsden, 2018, for another example of this). We draw on Plonsky and Oswald's (2014) Cohen's d field-specific benchmarks for interpreting our d values (within-subjects: 0.60 (small), 1.00 (medium), 1.40 (large); between-subjects: .40 (small), .70 (medium), 1.00 (large)), as well as ES izes for relevant interventions found by relevant meta-analyses (Shintani, Li & Ellis, 2013) and individual studies (Marsden, 2006; Marsden & Chen, 2011).

RESULTS

Habitual IMP in the Picture-Based Oral Narrative

A statistically significant two-way interaction between Time and Group ($F(5, 112) = 7.662, p = .000, \eta^2 = .275, \eta_p^2 = .264$) indicated between-group differences for appropriate IMP use over time. Statistically significant main effects for Time ($F(1.8, 112.1) = 43.705, p = .000, \eta^2 = .505, \eta_p^2 = .406$) and Group ($F(3,64) = 16.522, p = .000, \eta^2 = .220, \eta_p^2 = .436$) were also found.

Between-Group Differences in Habitual IMP Use. Group scores were compared at Pretest, Posttest, and Delayed Posttest (see Table 3).

At Pretest, comparisons confirmed no between-group differences (all CIs for d included zero, see Table 3). Appropriate IMP use for habitual events ranged from 31%-36% across all groups (see Table 3). Other forms inappropriately used in these past habitual contexts included *Passé Composé* (35%, examples 1-3) and, to a lesser extent, PRES (18%, examples 4-6).

1. *pendant sa jeunesse chaque soir (erm) Alex (erm) a fait erm ses devoirs (participant 214)*

‘during her youth, every evening Alex (erm) did^{-PAST PERFECTIVE} (erm) her homework’

2. *donc chaque matin Nathalie a lu son livre préféré à ses poupées (participant 219)*

‘so every morning Natahalie read^{-PAST PERFECTIVE} her favourite book to her dolls’

3. *pendant sa jeunesse chaque soir Alex elle a écrit beaucoup (participant 228)*

‘during her youth, every evening Alex wrote^{-PAST PERFECTIVE} a lot’

4. *chaque soir pendant sa jeunesse Alex fait des choses très calme (participant 212)*

‘every evening during her youth Alex does^{-PRESENT} things very calmly’

5. *chaque matin Nathalie peint un image et construit un maison des boîtes (participant 224)*

‘every morning Nathalie paints^{-PRESENT} a picture and builds^{-PRESENT} a house out of boxes’

6. *pour Pompon le chat (erm) chaque matin il dort (participant 242)*

‘for Pompon the cat (erm) every morning he sleeps^{-PRESENT}

Following training at Posttest, comparisons with Control showed large differences because of more appropriate IMP use in the treatment groups. At Delayed Posttest, only the L2+L1 group’s use of IMP to express past habituality was more appropriate than Control (large ES). We found no differences between (a) Control and L2+L1prac (negligible ES) and (b) Control and L2-only (negligible ES).

Two of the between-treatment-group comparisons at Posttest showed small but unreliable differences: L2+L1’s use of IMP was slightly more appropriate than L2+L1prac (small but unreliable ES because CIs for *d* included zero); L2+L1 and L2-only performed similarly (negligible ES). IMP scores in the L2-only group were higher than L2+L1prac (medium ES).

At Delayed Posttest, L2+L1’s scores were higher than both L2+L1prac (large ES) and L2-only (large ES). There were no Delayed Posttest differences between L2-only and L2+L1prac (negligible ES).

< TABLE 3 HERE >

< TABLE 4 HERE >

TABLE 3
Means (and Standard Deviations) for Habitual IMP (%TLU) in the Picture-Based Oral Narrative

	Pretest	Posttest	Delayed Posttest
L2+L1 (<i>n</i> =17)	31.18 (21.13)	80.51 (14.46)	76.10 (13.12)
L2+L1prac (<i>n</i> =19)	36.55 (22.75)	73.15 (7.58)	46.57 (24.92)
L2-only (<i>n</i> =17)	36.58 (21.61)	82.29 (11.8)	43.83 (22.19)
Control (<i>n</i> =16)	35.33 (23.24)	36.63 (23.54)	40.30 (24.52)

TABLE 4
Between-Group Comparisons for Habitual IMP in the Picture-Based Oral Narrative at Each Test Point (Mean Difference, Mean Standard Error (SE), *p*, and Cohen's *d* Effect Size [with CIs for *d*])

	Pretest		Posttest		Delayed Posttest	
	Mean difference (SE)	<i>p</i> , <i>d</i> [CIs]	Mean difference (SE)	<i>p</i> , <i>d</i> [CIs]	<i>p</i> , <i>d</i> [CIs]	<i>p</i> , <i>d</i> [CIs]
L2+L1 vs. L2+L1prac	-5.37 (7.31)	.883, -.24 [-.90, .42]	7.36 (3.91)	.264, .65 [-.04, 1.30]	29.54 (6.68)	.001, 1.46 [.69, 2.16]
L2+L1 vs. L2-only	-5.40 (7.33)	.882, -.25 [-.92, .43]	-1.79 (4.53)	.979, -.13 [-.80, .54]	32.28 (6.25)	.000, 1.77 [.94, 2.52]
L2+L1 vs. Control	-4.15 (7.75)	.950, -.19 [-.87, .50]	43.87 (6.85)	.000, 2.26 [1.34, 3.07]	35.81 (6.91)	.000, 1.84 [.98, 2.60]
L2-only vs. L2+L1prac	-.03 (7.39)	1.00, .00 [-.66, .65]	-9.14 (3.35)	.051, .93 [.22, 1.60]	2.74 (7.97)	.986, -.12 [-.77, .54]
L2+L1prac vs. Control	1.22 (7.81)	.999, .05 [-.61, .72]	36.52 (6.14)	.000, 2.17 [1.29, 2.95]	6.27 (8.49)	.881, .25 [-.42, .92]
L2-only vs. Control	1.25 (7.82)	.999, .06 [-.63, .74]	45.66 (6.54)	.000, 2.48 [1.52, 3.31]	3.53 (8.16)	.972, .15 [-.54, .83]

Note. Shading indicates reliable and meaningful ES because CIs for *d* do not include zero. The order of the groups in the first column can be used to interpret the direction of the ES. For example, group x vs group y, would show a positive ES if x outperformed y, but a negative ES if y > x.

Within-Group Changes in Habitual IMP Use. We compared performance between the three test points (see Table 5). In the Control group, no reliable changes were found over time (negligible ES). All treatment groups improved between Pretest and Posttest (large ES). However, between Posttest and Delayed Posttest, appropriate IMP use decreased majorly for both L2+L1prac (large ES) and L2-only (large ES), to the extent that Pretest-Delayed scores were not different (negligible ES). In contrast, we found no differences between L2+L1's Posttest and Delayed Posttest scores (negligible ES), indicating that their Pretest-Posttest improvement was maintained.

Parallel coordinate plots (see Figure 2) show these trajectories in detail (each line represents an individual learner), indicating detectable improvement between Pretest and Posttest for almost all individuals in the treatment groups. These Pre-Post improvement trajectories largely disappear for individuals in the L2-only and L2+L1prac groups, but not for those in the L2+L1 group. Individual performance in the Control group, however, is varied, without any discernible patterns over time.

Taken together, these results suggest that all three interventions improved learners' appropriate habitual IMP use in semi-spontaneous oral production immediately after instruction (i.e., at Posttest). However, these gains were maintained six weeks later only for learners who had received L1 EI (i.e., the L2+L1 group).

<TABLE 5 HERE>

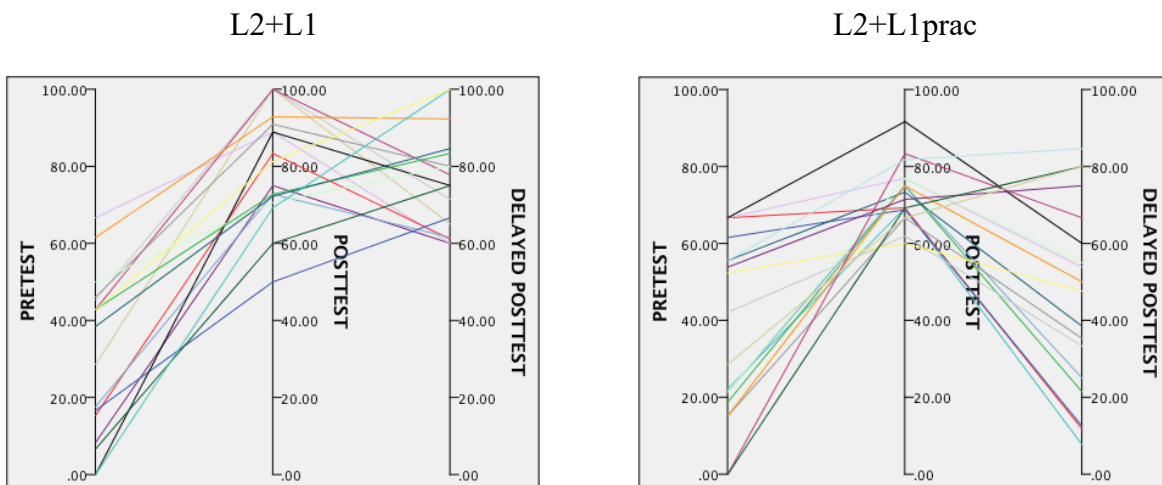
< FIGURE 2 HERE>

TABLE 5
 Within-Group Comparisons for Habitual IMP in the Picture-Based Oral Narrative (Mean Difference, Mean Standard Error (SE), p , and Cohen's d Effect Size with CIs for d)

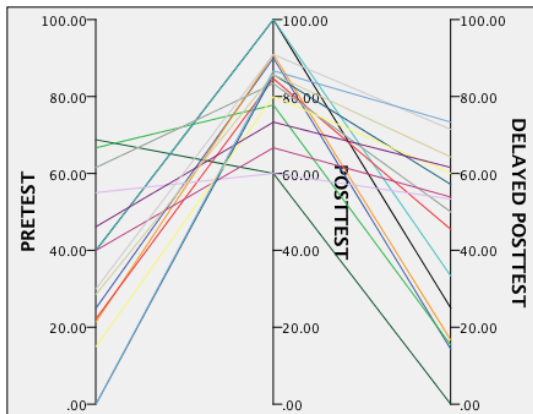
	Pretest vs. Posttest		Pretest vs. Delayed Posttest		Posttest vs. Delayed Posttest	
	Mean difference (SE)	p, d [CIs]	Mean difference (SE)	p, d [CIs]	Mean difference (SE)	p, d [CIs]
L2+L1 ($n=17$)	-49.32 (4.75)	.000, 2.72 [1.74, 3.58]	-44.92 (5.58)	.000, 2.55 [1.60, 3.39]	4.40 (4.83)	.376, -.32 [-.99, .36]
L2+L1prac ($n=19$)	-36.60 (5.42)	.000, 2.16 [1.32, 2.91]	-9.70 (8.49)	.270, .42 [-.23, 1.05]	26.33 (5.59)	.000, -1.44 [-2.12, -.070]
L2-only ($n=17$)	-45.71 (6.99)	.000, 2.63 [1.66, 3.47]	-7.24 (8.95)	.430, .33 [-.35, 1.00]	38.46 (6.06)	.000, -2.16 [-2.95, -1.27]
Control ($n=16$)	-1.30 (8.42)	.879, .06 [-.64, .75]	-4.97 (10.22)	.634, .21 [-.49, .90]	-3.67 (8.89)	.686, .15 [-.55, .84]

Note. Shading indicates reliable and meaningful ES because CIs for d do not include zero. The order of the groups in the first column can be used to interpret the direction of the ES. For example, group x vs group y , would show a positive ES if x outperformed y , but a negative ES if $y > x$.

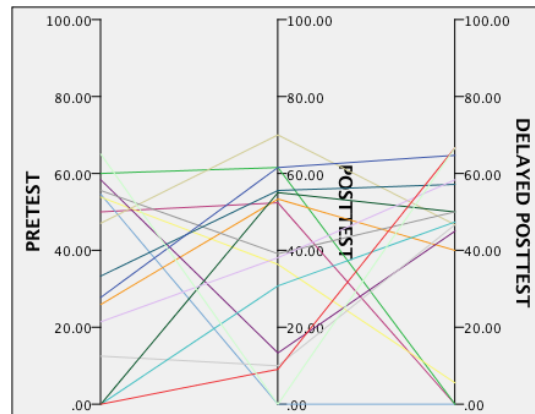
FIGURE 2
 Parallel Coordinate Plots of Habitual IMP in the Picture-Based Oral Narrative



L2-only



Control



Ongoing IMP in the Activity Description Oral Production Test

A statistically significant two-way interaction between Group and Time ($F(4, 97) = 9.285, p = .000, \eta^2 = .176, \eta_p^2 = .300$) indicated that ongoing IMP use varied between groups as a function of test point. There were also statistically significant main effects for Group ($F(3,65) = 33.957, p = .000, \eta^2 = .323, \eta_p^2 = .610$) and Time ($F(1.5, 97) = 83.680, p = .000, \eta^2 = .501, \eta_p^2 = .563$).

Between-Group Differences in Ongoing IMP Use. See Table 7 for all between-group comparisons. At Pretest, there were no meaningful between-group differences (all CIs for d passed through zero). Scores ranged from 35%-40% across all groups (see Table 6). Other forms inappropriately used in these past ongoing contexts included PRES (30%, examples 7-9) and, to a lesser extent, auxiliary + infinitive / present participle invented forms (16%, examples 10-12).

7. *il quitte son travail (participant 219)*

'he leaves^{-PRESENT} his job

8. *il sonne la cloche (participant 206)*

'he rings^{-PRESENT} the bell'

9. *elle regarde un film (participant 250)*

'she watches^{-PRESENT} a film'

10. *il était écrivant un lettre (participant 228)*

'he was^{-AUXILIARY-PAST} writing^{-PRESENT PARTICIPLE} a letter

11. *il était sonner la cloche (participant 247)*

'he was^{-AUXILIARY-PAST} ringing^{-INFINITIVE} the bell

12. *il était faisant le ski (participant 242)*

'he was^{-AUXILIARY-PAST} skiing^{-PRESENT PARTICIPLE}'

At both Posttest and Delayed Posttest, all treatment groups' IMP use was more appropriate than the Control group (large ES for all treatment group vs. control comparisons). These results contrast with our findings for habitual IMP, which showed no between-group differences at Delayed Posttest between (a) Control and L2+L1prac and (b) Control and L2-only.

Comparisons between the treatment groups showed no reliable differences at Posttest or Delayed Posttest. At Posttest, comparisons between L2+L1 versus L2+L1prac revealed a small but unreliable difference (CIs for *d* included zero) due to slightly higher scores in the L2+L1 group. No differences were found between L2+L1 and L2-only (negligible ES) and L2+L1prac and L2-only (negligible ES). At Delayed Posttest, no differences were found between L2+L1 versus L2+L1prac (negligible ES) and L2+L1 and L2-only (negligible ES). A small but unreliable difference (CIs for *d* included zero) was found between L2+L1prac and L2-only due to slightly higher scores in the L2-only group.

<TABLE 6 HERE>

<TABLE 7 HERE>

<TABLE 8 HERE>

TABLE 6
Means (and Standard Deviations) for Ongoing IMP in the Activity Description Oral Production Test

	Pretest <i>M(SD)</i>	Posttest <i>M(SD)</i>	Delayed Posttest <i>M(SD)</i>
L2+L1 (<i>n</i> =17)	35.95 (21.17)	80.66 (7.66)	77.15 (14.09)
L2+L1prac (<i>n</i> =19)	36.19 (22.32)	76.14 (9.26)	73.14 (9.14)
L2-only (<i>n</i> =17)	40.81 (17.74)	79.29 (9.08)	77.88 (10.33)
Control (<i>n</i> =16)	38.27 (21.59)	34.26 (18.93)	40.83 (19.68)

TABLE 7
Between-Group Comparisons for Ongoing IMP in Activity Description Oral Production Test at Each Test Point (Mean Difference, Mean Standard Error (SE), *p*, and Cohen's *d* Effect Size with CIs for *d*)

	Pretest		Posttest		Delayed Posttest	
	Mean difference (SE)	<i>p</i> , <i>d</i> [CIs]	Mean difference (SE)	<i>p</i> , <i>d</i> [CIs]	Mean difference (SE)	<i>p</i> , <i>d</i> [CIs]
L2+L1 vs. L2+L1prac	-0.24 (7.25)	1.00, -.01 [-.67, .64]	4.52 (2.82)	.391, .53 [-.15, 1.18]	3.55 (4.01)	.813, .34 [-.32, .99]
L2+L1 vs. L2-only	-4.86 (6.69)	.886, -.25 [-.92, .43]	1.37 (2.88)	.964, .16 [-.51, .83]	-.72 (4.24)	.998, -.06 [-.73, .61]
L2+L1 vs. Control	-2.32 (7.45)	.989, -.11 [-.79, .58]	46.40 (5.08)	.000, 3.25 [2.15, 4.20]	36.32 (5.99)	.000, 2.13 [1.23, 2.93]
L2-only vs. L2+L1prac	-4.62 (6.69)	.900, .23 [-.43, .88]	-3.14 (3.06)	.734, .34 [-.32, .99]	-4.27 (3.27)	.565, .49 [-.19, 1.14]
L2+L1prac vs. Control	-2.08 (7.44)	.964, -.09 [-.76, .57]	41.88 (5.19)	.000, 2.89 [1.89, 3.76]	32.77 (5.35)	.000, 2.17 [1.29, 2.95]
L2-only vs. Control	2.54 (6.90)	.983, .13 [-.56, .81]	45.03 (5.22)	.000, 3.06 [2.00, 3.98]	37.05 (5.52)	.000, 2.38 [1.44, 3.20]

Note. Shading indicates reliable and meaningful ES because CIs for *d* do not include zero. The order of the groups in the first column can be used to interpret the direction of the ES. For example, group x vs group y, would show a positive ES if x outperformed y, but a negative ES if $y > x$.

TABLE 8.
 Within-Group Comparisons for Ongoing IMP in the Description Oral Production Test (Mean Difference, Mean Standard Error (SE), p , and Cohen’s d Effect Size with CIs for d)

	Pretest vs. Posttest		Pretest vs. Delayed Posttest		Posttest vs. Delayed Posttest	
	Mean difference (SE)	p, d [CIs]	Mean difference (SE)	p, d [CIs]	Mean difference (SE)	p, d [CIs]
L2+L1 ($n=17$)	-44.71 (5.67)	.000, 2.81 [1.81, 3.68]	-41.19 (7.05)	.000, 2.29 [1.38, 3.09]	3.51 (3.04)	.265, -.31 [-.98, .37]
L2+L1prac ($n=19$)	-39.95 (5.73)	.000, 2.34 [1.47, 3.11]	-37.41 (5.66)	.000, 2.17 [1.33, 2.92]	2.54 (1.97)	.214, -.33 [-.96, .32]
L2-only ($n=17$)	-38.48 (4.79)	.000, 2.73 [1.74, 3.59]	-37.07 (5.02)	.000, 2.55 [1.60, 3.39]	1.41 (2.32)	.551, -.14 [-.81, .53]
Control ($n=16$)	4.01 (6.59)	.552, -.20 [-.89, .50]	-2.56 (8.19)	.759, .12 [-.57, .81]	-6.57 (6.29)	.313, .34 [-.37, 1.03]

Note. Grey shading indicates reliable and meaningful ES because CIs for d do not include zero. The order of the groups in the first column can be used to interpret the direction of the ES. For example, group x vs group y, would show a positive ES if x outperformed y, but a negative ES if $y > x$.

Within-Group Changes in Ongoing IMP Use Over Time. See Table 8 for all within-group comparisons. For the Control group, scores did not change over time (negligible ES for all comparisons). For all treatment groups, we found major improvement between Pretest and Posttest (large ES) and between Pretest and Delayed Posttest (large ES). There was no reliable change for any treatment group between Posttest-Delayed Posttest (negligible ES).

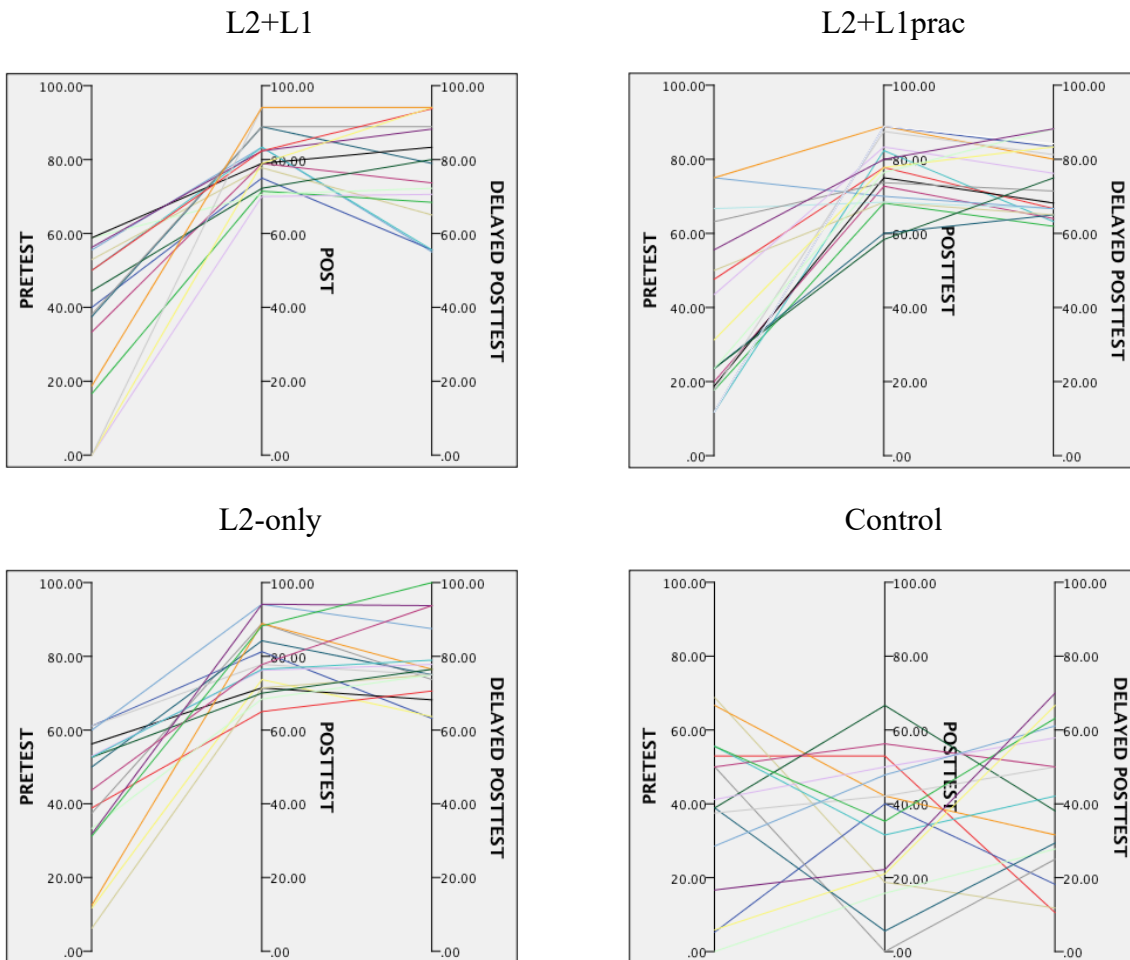
The parallel coordinate plots in Figure 3 show these trajectories at the level of individual learners, showing detectable improvement for almost all individuals in the treatment groups. Performance in the Control group, however, is varied. These visualizations show a remarkably clear and consistent effect of instruction across individuals’ performances.

In sum, our results indicate that the L2+L1, L2+L1prac, and L2-only treatments all led to more appropriate use of both habitual and ongoing IMP immediately after instruction (i.e., at Posttest). However, different patterns of results were found at Delayed Posttest (six weeks later): For habitual IMP, only the L2+L1 group retained their gains at Delayed Posttest; for ongoing IMP, all treatment groups retained their gains.

<FIGURE 3 HERE>

FIGURE 3

Parallel Coordinate Plots of Ongoing IMP in the Activity Description Oral Production Test



DISCUSSION

The present study examined the extent to which different types of EI about viewpoint aspect (L2 only vs. L2 + L1) and comprehension practice (of L2 sentences only vs. of L2 + L1 sentences) improved L2 learners' oral production of the French IMP, immediately after instruction (at Posttest) and then six weeks later (at Delayed Posttest).

All treatments improved learners' habitual IMP use in oral production in a discourse-level test immediately after the instruction, but six weeks later only the effects of L2+L1 treatment - the only treatment that included EI about the L1 - were detectable. For past ongoing events, we found major improvement for all treatments between Pretest-Posttest (large ES), and these gains were retained at Delayed Posttest (negligible ES between Posttest and Delayed Posttest). In sum, all treatments appeared to improve learners' use of *ongoing* IMP in oral production immediately after the instruction with effects additionally detectable six weeks later, but *only* the L2+L1 treatment improved *habitual* IMP in ways that were still observable six weeks later.

These oral production results are consistent with McManus and Marsden's (2017, 2018) previously discussed findings for comprehension, which showed that the L2+L1 treatment (i.e., providing L1 EI with L1 practice alongside a core of L2 EI with L2 practice) improved the speed (self-paced reading test) and accuracy (sentence judgement test in reading and listening) of L2 comprehension of habitual and ongoing IMP immediately after instruction with gains retained six weeks later. As for the L2+L1prac and L2-only treatments, however, we found marginally more accurate performance in the oral production tasks than in the comprehension tasks at immediate Posttest. It is likely that differences in the nature of the tasks could explain why learners appeared to perform better in oral production than in comprehension. First, the comprehension tests required learners to respond to specific uses of IMP in pre-determined

sentences, while there was, to a certain extent, more flexibility in the production tests for learners to use a variety of linguistic resources to express particular viewpoint aspect meanings. Second, the Picture-Based Narrative, which elicited habitual IMP, was a discourse level task that required learners to narrative a story, whereas the Activity Description Oral Production Task, which elicited ongoing IMP, was more mechanical and provided learners with verbs to use in sentences. In many respects, the Activity Description Oral Production Task was less demanding than Picture-Based Narrative. These could be possible explanations for why the L2+L1prac and L2-only appeared to perform better with ongoing IMP than habitual IMP at Delayed Posttest in oral production than in comprehension.

Taken together, then, two trends emerge from the current study's oral production findings and those for comprehension as reported in McManus and Marsden (2017, 2018). First, at immediate Posttest, all treatments improved their oral production of ongoing and habitual IMP, but only the L2+L1 and (to a lesser extent) L2-only treatments improved comprehension. Second, at Delayed Posttest, only the L2+L1 treatment led to improved production and comprehension of *both* ongoing *and* habitual IMP. Thus, our findings indicate that oral production and comprehension improvement for habitual IMP was only found to be detectable six weeks after instruction for learners whose treatment included L1 EI, combined with L1 practice and the core, L2 EI and practice.

Our findings also enrich those of McManus & Marsden (2019), that found that automaticity (i.e., less variability in speed as accurate responses got faster) was more likely to be evidenced during the comprehension practice itself in the group receiving the EI about the L1 compared to the other groups. Our current findings suggest that this during-practice 'automaticity' benefit is likely to have contributed to the gains observed after practice in the oral production tests, at least for use of habitual IMP.

An important finding that requires explanation is why EI about L2 was sufficient to improve learners' use of IMP's ongoing function, but not its habitual function. Only additional EI about L1 improved the accuracy of habitual IMP use. Different cue validities in L1 for ongoingness versus habituality could explain these findings.

L1 Explicit Instruction to Address Low Cue Validity in L1

As previously discussed, SLA research on the acquisition of polyfunctional aspectual forms (like IMP) has shown that a form's different functions tend to be acquired in stages, rather than all at once (Andersen, 1984; Bardovi-Harlig, 2000; Salaberry, 2008). Explanations for the acquisition order of these functions, as well more general explanations about L2 learnability problems, have tended to focus on variations in the availability and reliability of cues in the L2 and learners' (in)attention to them (Andersen and Shirai, 1994; Ellis & Sagarra, 2011; Schmidt, 1990; Zhao & MacWhinney 2018). However, very little research has considered low cue validities in L1 as potential explanations for L2 learnability problems, even though many theories of SLA do forefront critical roles for L1 knowledge in L2 learning (e.g., Ellis 2006; O'Grady 2013; MacWhinney, 2012). For example, Zhao and MacWhinney (2018) proposed that variations in the availability and reliability of English (i.e., the L2) cues for (in)definiteness can explain Mandarin Chinese speakers' difficulties learning English articles; in addition, the low availability and reliability of Mandarin Chinese (i.e., the L1) cues for (in)definiteness could be a further explanation for this learnability problem (see Chen 2015).

This is one likely explanation for why IMP's habitual function appears later acquired than its ongoing function among English L1 speakers: the low validity of English cues for habituality reduces learners' sensitivity/attention to the concept of habituality itself, which, in turn, delays learning of L2 cues indexing that meaning (Athanasopoulos & Bylund, 2013;

MacWhinney, 2012). This explanation is borne out in our results because only the group receiving EI about L1 cues for habituality (L2+L1 group) demonstrated L2 learning of IMP's habitual function leading to knowledge that was available for use in an oral production test (as evidenced by Delayed Posttest performance). This L1 EI was designed to increase learners' sensitivity to (a) the concept of habituality and (b) L1 cues for habituality. The L2 EI, in contrast, was insufficient for learning IMP's habitual function for oral production, arguably because the L2 EI only focused on L2 cues for habituality, which did not address the nature of the learning problem in a sufficiently explicit manner. However, the L2 EI (received by all treatment groups) was sufficient for learning IMP's *ongoing* function, probably given the relative conceptual saliency of ongoingness to these speakers due to high cue validity in L1 for this meaning. In sum, these results suggest that L1 EI was necessary for learning IMP's habitual function because of English speakers' reduced awareness of this concept (a consequence of the low validity of English cues indexing habituality). The L1 EI benefitted performance by increasing learners' awareness of (a) the concept of habituality and (b) L1 cues for habituality, which better allowed mapping of L2 cues to the concept of habituality and inhibiting (or transferring) use of L1 cues.

Therefore, in addition to different cue validities between L1 and L2 (i.e. the extent to which the same cues index the same meanings in L1 and L2), L1 cue validities are argued to play an important role in understanding L2 development. This is because there is likely to be reduced sensitivity when a concept is indexed by a variety of cues in the L1. Such learning situations may benefit from EI about L1 to increase awareness of low cue validities in L1.

Our finding that additional L1 practice (i.e., interpretation of English sentences), when not accompanied by EI about the L1, did not benefit IMP's habitual function supports this conclusion: practice interpreting L1 habitual cues without EI about these cues was insufficient to increase sensitivity to the low validity of L1 cues for the concept of past habituality. Therefore,

in addition to characteristics of L2 cues (e.g., availability, reliability) and different cue validities in L1 and L2 (which are already demonstrated to take on critical roles in SLA), L1 cue validities should also be considered important for understanding L2 development. EI about L1 can facilitate L2 learning by increasing learners' awareness of low L1 cue validities.

Limitations and Future Research

Due to the small number of participants in each group, we note that our findings are tentative. We also note that we did not elicit the IMP's habitual and ongoing functions in a single test, but instead used different tests for each function. For these reasons, our conclusions require replication. The habitual test was a semi-spontaneous, discourse-level oral production test which required learners to construct a narrative, whereas the ongoing test was more controlled and mechanical in order to set up contexts to elicit ongoingness. It is possible that performance was less demanding in the ongoing test and allowed (more) access to a more explicit knowledge type. However, we note that no change was found for the Control group, thus weakening the likelihood that artefacts of test design are entirely responsible for our findings. If test type alone explained our findings, then the Control group could have drawn on existing EI about L2 past ongoingness, which is certainly part of their formal curriculum prior to the current study, and, as evidenced in baseline scores, almost all the participants across all groups did indicate some existing knowledge of the ongoing use of IMP. Given the lack of gains in the Control group, we consider it unlikely that the ongoing test simply allowed gains to be observed more easily. We also note, as previously discussed, that previous empirical and theoretical SLA research corroborates the notion that IMP's ongoing function is more easily acquired (and therefore likely to be more sensitive to instruction) than the habitual function by English speakers, providing secondary support for our claims.

Notwithstanding these limitations, our findings provide a number of directions for future research on differences between instructional components and their impact on L2 learning. For example, it is unclear whether systematic production practice (L2 vs L2+L1), instead of comprehension practice, would lead to the same learning gains, or the extent to which altering the amount or spacing of practice would affect the findings (see Kasprovicz & Marsden, under review; Suzuki, 2017). As previously noted, learners completed extensive L2 practice, but very little L1 practice in comparison. Although additional L1 practice without L1 EI (the L2+L1prac group) appeared to provide few additional learning benefits, larger amounts of L1 practice may lead to different results. Also, future research might even explore the effects of providing *only* L1 EI and L1 practice (i.e., without L2 treatments) for features with L1-L2 form-meaning differences so as to isolate the effects of clarifying L1 form-meaning mappings for L2 learning, especially perhaps in contexts, such as with advanced learners, where some use of the L2 forms is already established.

In addition, future research should investigate potential interactions between proficiency and instructional effectiveness by studying the outcomes of instruction among learners with different amounts/types of language exposure and/or L2 proficiency. For example, Isabelli (2007) found that instruction about Spanish Subjunctive was more effective for learners who had recently returned from study abroad than for learners who had not studied abroad, indicating potential interactions between language exposure and/or proficiency and instructional effectiveness. Since the current study did not investigate such factors (as all the participants were advanced, classroom learners), it remains an empirical question whether the same patterning of results would be found for less experienced learners or for learners with less classroom experience but more language exposure (e.g., following study abroad, as in Isabelli 2007).

CONCLUSION

The current study examined the extent to which differences in the type of EI and comprehension practice improved the appropriacy of IMP use in L2 oral production. We provided three comprehension-based treatments: one group received EI about the L2 plus extensive L2 comprehension practice (L2-only group); a second group received the same L2 EI, L2 comprehension practice, plus additional L1 comprehension practice (L2+L1prac group); and a third group received the same L2 EI, L2 comprehension practice, L1 comprehension practice, plus additional EI about the L1 (L2+L1 group). A Control group received no instruction and completed only the Pretest, Posttest, and Delayed Posttest. This design allowed us to examine how differences in the type of EI (about the L2 vs. about the L2+L1) and type of comprehension practice (L2 only vs. L2+L1) impacted L2 learning of viewpoint aspect in L2 French. Compared to L2-only and L2+L1prac, results showed that providing additional L1 EI benefitted the oral production of *both* habitual and ongoing IMP at six weeks after treatment. The L2-only and L2+L1prac treatment groups made gains at Posttest for both IMP meanings, but these were only maintained at Delayed for ongoing IMP. For habitual IMP, providing EI about the L1 provided more lasting benefits than the other treatments.

Taken together, we argue that the low validity of L1 English cues for habituality reduced English speaker learners' sensitivity to this concept and the cues that index it. L1 EI was needed to improve the L2 learning of habitual IMP because it helped concretize a concept of past habituality that was more useful, to them as L1 English speakers, for learning French IMP. We suggest that this helped learners to work out complex relations between L1-L2 form-meaning mappings, hypothesized to be a cause of L2 learning difficulty. Since, compared to habituality, ongoingness has a relatively less complex L1 cue system and is expressed morphologically, by one reliable cue, in both the L1 and L2, additional EI about the L1 appeared to provide no extra

learning benefits for oral production. These results suggest that tailoring instruction, specifically the nature of the EI, to the nature of the learning problem can facilitate L2 learning. In particular, EI about L1 can facilitate L2 learning by increasing learners' awareness of low L1 cue validities.

NOTES

1. All treatment groups spent the same amount of time on the L2 EI and L2 practice. Although the additional L1 EI and L1 practice components slightly extended the length of the treatments for the L2+L1 and L2+L1prac groups, these additions did not introduce major time differences between the treatments because the L1 EI was short and the L1 practice was provided in small amounts. See description of L1 EI and L1 practice for more information.
2. Based on a meta-analysis of reliability coefficients in L2 research, Plonsky and Derrick (2016) propose that .83 (median = .92) should be considered a general (not absolute) threshold for an acceptable estimate of interrater reliability.

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.
- Andersen, R. W. (1984). The one to one principle of interlanguage construction. *Language Learning*, 34, 77-95.

- Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition*, 16, 133-156.
- Andringa, S., de Glopper, K. & Hacquebord, H. (2011). Effect of explicit and implicit instruction on free written response task performance. *Language Learning*, 61, 868-903
- Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science*, 37, 286–309.
- Avery, N. & Marsden, E. (under review) A meta-analysis of sensitivity to grammatical information during self-paced reading in a second language. *Studies in Second Language Acquisition*.
- Ayoun, D. (2004). The effectiveness of written recasts in the second language acquisition of aspectual distinctions in French: A follow-up study. *The Modern Language Journal*, 88, 31-55.
- Ayoun, D. (2013). *The second language acquisition of French tense, aspect, mood and modality*. Amsterdam: John Benjamins
- Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell.
- Bartning, I. (1997). L'apprenant dit avancé et son acquisition d'une langue étrangère. Tour d'horizon et esquisse d'une caractérisation de la variété avancée [The advanced learner and their acquisition of a foreign language. Overview and outline of the advanced learner variety]. *Acquisition et interaction en langue étrangère*, 9, 9–50.
- 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.

- Calabria, M., Costa, A., Green, D. W., & Abutalebi, J. (2018). Neural basis of bilingual language control. *Annals of the New York Academy of Sciences*, 1426, 221-235.
- 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.
- Dahl, Ö., & Velupillai, V. (2013). Perfective/imperfective aspect. In Dryer, M. S. & Haspelmath, M. (eds). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed 13 January 2019 at <http://wals.info/chapter/65>
- DeKeyser, R. (2017). Knowledge and skill in ISLA. In S. Loewen and M. Sato (Eds.), *Routledge Handbook of Instructed Second Language Acquisition* (pp. 15–32). London: Routledge.
- DeKeyser, R. M., & Sokalski, K. J. (1996). The differential role of comprehension and production practice. *Language Learning*, 46, 613–642.
- Domínguez, L., Tracy-Ventura, N., Arche, M. J., Mitchell, R., & Myles, F. (2013). The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology. *Bilingualism: Language and Cognition*, 16, 558–577.
- Ellis, N. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27, 164-194.
- Ellis, N. C. (2008). The associative learning of constructions, learned attention, and the limited L2 endstate. In Robinson, P. & Ellis, N.C. (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372-405). New York, NY: Routledge

- Ellis, N. C., & Sagarra, N. (2011). Learned attention in adult language acquisition: A replication and generalization study and meta-analysis. *Studies in Second Language Acquisition*, 33, 589-624.
- Ellis, N. C., Hafeez, K., Martin, K. I., Chen, L., Boland, J., & Sagarra, N. (2014). An eye-tracking study of learned attention in second language acquisition. *Applied Psycholinguistics*, 35, 547-579.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. London: Routledge.
- Fernández, C. (2008). Reexamining the role of explicit information in processing instruction. *Studies in Second Language Acquisition*, 30, 277-305.
- Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics*. Los Angeles: Sage.
- Henry, N., Jackson, C. N., & DiMidio, J. (2017). The role of prosody and explicit instruction in processing instruction. *The Modern Language Journal*, 101, 294–319.
- 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, 331–349.
- Howard, M. (2005). Les contextes prototypiques et marqués de l'emploi de l'imparfait par l'apprenant du français langue étrangère [The prototypical and marked contexts of Imperfect use by the advanced learner of French]. In E. Labeau & P. Larrivée (Eds.), *Nouveaux développements de l'imparfait* [New developments of the Imperfect] (pp. 175–197). New York, NY: Rodopi.

- Isabelli, C. A. (2007). Development of the Spanish subjunctive by advanced learners: Study abroad followed by at-home instruction. *Foreign Language Annals*, 40, 330-341.
- Isabelli, C. A. (2008). First noun principle or L1 transfer principle in SLA?. *Hispania*, 465-478.
- 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, 350-368.
- Kasprowicz, R., & Marsden, E. (2018). Towards ecological validity in research into input-based practice: Form spotting can be as beneficial as form-meaning practice. *Applied Linguistics*.
- Kasprowicz, R. , & Marsden, E. (under review). Distribution of practice effects for the learning of foreign language verb morphology in the young learner classroom. *The Modern Language Journal*.
- Kihlstedt, M. (1998). *La référence au passé dans le dialogue: Etude de l'acquisition de la temporalité chez des apprenants dits avancés de français [Reference to the past in dialogue: A study of the acquisition of temporality among advanced learners of French]*. 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, 74–106.
- 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.

- Larsen-Freeman, D. (2015). Research into practice: Grammar learning and teaching. *Language Teaching*, 48, 263-280.
- Larson-Hall, J. (2014). *A guide to doing statistics in second language research using SPSS and R*. London: Routledge.
- Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, 65, 127-159.
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied linguistics*, 29, 694-716.
- Li, M., & DeKeyser, R. (2017). Perception practice, production practice, and musical ability in L2 Mandarin tone-word learning. *Studies in Second Language Acquisition*, 39, 593-620.
- Lightbown, P. M. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Han, Z. & Park, E.S. (Eds.), *Understanding second language process* (pp. 27-44). Bristol, UK: Multilingual Matters
- Lonsdale, D., & Le Bras, Y. (2009). *A frequency dictionary of French: Core vocabulary for learners*. London: Routledge.
- Lorenzo, C. (2002). *Les relations temporo-aspectuelles dans le récit oral en français et en castillan, langues premières et langues étrangères: étude transversale du stade ultime de l'acquisition d'une langue étrangère [Tempo-aspectual relations in oral retellings in French and Castilian as first and second languages: A crosslinguistic study of the end state of foreign language acquisition]*. (Unpublished doctoral dissertation). Nanterre: Université de Nanterre – Paris X.
- 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

- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk: The database*. Mahwah, NJ: Lawrence Erlbaum Associates.
- 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, NY: Routledge.
- Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, 6, 97-115.
- Marsden, E. (2006). Exploring input processing in the classroom: An experimental comparison of processing instruction and enriched input. *Language Learning*, 56, 507-566.
- Marsden, E., & Chen, H. Y. (2011). The roles of structured input activities in processing instruction and the kinds of knowledge they promote. *Language Learning*, 61, 1058-1098.
- Marsden, E., Myles, F., Rule, S., & Mitchell, R. (2003). Using CHILDES tools for researching second language acquisition. In Sarangi, S. & van Leeuwen, T. (eds). *British Studies in Applied Linguistics* (pp. 98-113). London: Continuum.
- McManus, K. (2013). Prototypical influence in second language acquisition: What now for the Aspect Hypothesis? *International Review of Applied Linguistics in Language Teaching*, 51, 299-322.
- McManus, K. (2015). L1-L2 differences in the acquisition of form-meaning pairings in a second language. *Canadian Modern Language Review*, 71, 155-181.
- McManus, K., & Marsden, E. (2017). L1 explicit instruction can improve L2 online and offline performance. *Studies in Second Language Acquisition*, 39, 459-492.

- 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, 459-475.
- McManus, K., & Marsden, E. (2019). Signatures of automaticity during practice: Explicit instruction about L1 processing routines can improve L2 grammatical processing. *Applied Psycholinguistics*, 40, 205-234.
- Mitchell, R., Tracy-Ventura, N., & McManus, K. (2017). *Anglophone students abroad: Identity, social relationships, and language learning*. New York: Routledge.
- Murakami, A. (2016). Modeling systematicity and individuality in nonlinear second language development: The case of English grammatical morphemes. *Language Learning*, 66, 834-871.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, 50, 417-528.
- Norouzian, R., & Plonsky, L. (2018). Eta-and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*, 34, 257-271.
- O'Grady, W. (2013). The illusion of language acquisition. *Approaches to Bilingualism*, 3, 253-85.
- Parisse, C., & Le Normand, M. T. (2000). How children build their morphosyntax: The case of French. *Journal of Child Language*, 27, 267-292.
- Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition*, 6, 69-78.
- Plonsky, L. & Oswald, F. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning* 64, 878-912
- Ranta, L., & Lyster, R. (2017). Form-focused instruction. In Garrett, P. & Cotts, J. (Eds.), *The Routledge Handbook of Language Awareness* (pp. 58-74). New York, NY: Routledge

- Roberts, L., & Liszka, S. A. (2013). Processing tense/aspect-agreement violations on-line in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research*, 29, 413-439.
- Salaberry, R. M. (2008). *Marking past tense in second language acquisition: A theoretical model*. New York, NY: Continuum.
- Sanz, C., & Morgan-Short, K. (2004). Positive evidence versus explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning*, 54, 35-78.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied linguistics*, 11, 129-158.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime 2.0: User's guide*. Pittsburgh: Psychology Software.
- Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-Based versus Production-Based grammar instruction: A Meta-Analysis of comparative studies. *Language Learning*, 63, 296-329.
- 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
- Smith, C. (1997). *The parameter of aspect*. New York, NY: Springer:
- Spada, N., Lightbown, P., & White, J. (2005). The importance of form/meaning mappings in explicit form-focused instruction. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 199-234). Berlin: DeGruyter
- Starren, M. (2001). *The second time: The acquisition of temporality in Dutch and French as a second language*. Utrecht, NL: DeGruyter

- Suzuki, Y. (2017). The optimal distribution of practice for the acquisition of L2 morphology: A conceptual replication and extension. *Language Learning*, 67(3), 512-545.
- 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. & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning*, 64, 279–309
- VanPatten, B. (2017). Processing instruction. In S. Loewen and M. Sato (Eds.), *Routledge Handbook of Instructed Second Language Acquisition* (pp. 166–180). London: Routledge.
- VanPatten, B., & Oikkenon, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition*, 18, 495-510.
- VanPatten, B., Collopy, E., Price, J., Borst, S., & Qualin, A. (2013). Explicit information, grammatical sensitivity, and the first-noun principle: A cross-linguistic study in processing instruction. *The Modern Language Journal*, 97, 506–527.
- White, J. L., & Horst, M. (2012). Cognate awareness-raising in late childhood: Teachable and useful. *Language Awareness*, 21, 181–196.
- Wu, Y. J., & Thierry, G. (2010). Investigating bilingual processing: the neglected role of language processing contexts. *Frontiers in Psychology*, 1, 178.
- Wulff, S., & Ellis, N. C. (2018). Usage-based approaches to second language acquisition. In Miller, D., Bayram, F., Rothman, J. & Serratrice, L. (Eds.), *Bilingual Cognition and Language: The state of the science across its subfields* (pp. 37-56). Amsterdam: John Benjamins
- Zhao, H., & MacWhinney, B. (2018). The Instructed Learning of Form–Function Mappings in the English Article System. *The Modern Language Journal*, 102, 99-119.

APPENDIX

TABLE 1A

Description of the Core L2-Only Treatment (Received by all Treatment Groups) and the Additional L1 EI and Practice Used in Session 1: Ongoingness (Present vs Past). For all Materials, see McManus and Marsden (2017) and IRIS

	Core L2-only treatment	Additional L1 components								
Pre-practice EI	<p>[Watch a six-second video clip of man eating an apple. The apple was never fully eaten.]</p> <p>To describe this you could say: <i>Il mange une pomme</i> Or <i>Il mangeait une pomme</i></p> <p>The difference between these two is: Il mange = ongoing action RIGHT NOW Il mangeait = ongoing action IN THE PAST</p> <p>The ends of the verbs distinguish between an ongoing action in the present <i>versus</i> past e.g. [Four verbs presented in pairs, aurally and in writing]:</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="text-align: center;"><i>Présent</i></td> <td style="text-align: center;"><i>Imparfait</i></td> </tr> <tr> <td style="text-align: center;">RIGHT NOW</td> <td style="text-align: center;">IN PAST</td> </tr> <tr> <td style="text-align: center;">regarde</td> <td style="text-align: center;">regardait</td> </tr> <tr> <td style="text-align: center;">[Rəgard]</td> <td style="text-align: center;">[Rəgardɛ]</td> </tr> </table>	<i>Présent</i>	<i>Imparfait</i>	RIGHT NOW	IN PAST	regarde	regardait	[Rəgard]	[Rəgardɛ]	<p>[Same video as L2-only treatment]</p> <p>To describe this you could say: <i>He is eating an apple</i> Or <i>He was eating an apple</i></p> <p>The difference between these two is: ‘he is eating’ = ongoing action RIGHT NOW ‘he was eating’ = ongoing action IN PAST”</p> <p>To identify ongoing meaning in the present <i>versus</i> the past, you need to focus on the auxiliary.</p> <p>Look/listen out for ‘is’ or ‘was’ to indicate whether it is an ongoing action taking place RIGHT NOW (present) or it is one IN THE PAST (past).”</p>
<i>Présent</i>	<i>Imparfait</i>									
RIGHT NOW	IN PAST									
regarde	regardait									
[Rəgard]	[Rəgardɛ]									
Practice	<p>96 French items (48 listening, 48 reading).</p> <p>Aim: Identify whether an ongoing event is taking place:</p> <p style="text-align: center;">“MAINTENANT” (right now) or “DANS LE PASSÉ” (in the past)</p>	<p>Additional 32 English items (16 listening, 16 reading).</p> <p>Aim: identify whether an ongoing event is taking place:</p> <p style="text-align: center;">“RIGHT NOW” or “IN THE PAST”</p>								

Example (English glosses not provided): Example:

Il...

(1) fait du shopping ('is shopping')

(2) faisait du shopping ('was shopping')

He...

(1) is eating a sandwich

(2) was eating a sandwich

**EI given
immediately after
incorrect responses
during practice**

After incorrectly responding
'MAINTENANT':

"NOTE: The IMPARFAIT expresses an ongoing event DANS LE PASSÉ, not an ongoing event taking place MAINTENANT"

After incorrectly responding 'DANS LE PASSÉ':

"REMEMBER: The present tense in French expresses an ongoing event taking place MAINTENANT; not an ongoing action DANS LE PASSÉ"

After incorrectly responding 'RIGHT NOW':

"The present tense in English ('is +ing') and in French expresses the same meaning: ongoing action taking place RIGHT NOW"

After incorrectly responding 'IN THE PAST':

"The past tense in English ('was +ing') is the same as the IMP in French (-ait). They both express an ongoing action IN THE PAST"
