Preprint version
Applied Psycholinguistics, accepted for publication 09 July 2018
kmcmanus@psu.edu

SIGNATURES OF AUTOMATICITY DURING PRACTICE: EXPLICIT INSTRUCTION ABOUT L1 PROCESSING ROUTINES CAN IMPROVE L2 GRAMMATICAL PROCESSING

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

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

This study examined the extent to which explicit instruction about L1 and L2 processing routines improved the accuracy, speed, and automaticity of learners' responses during sentence interpretation practice. Fifty-three English-speaking learners of L2 French were assigned to one of the following treatments: (1) a 'core' treatment consisting of L2 explicit information (EI) with L2 interpretation practice (L2-only group), (2) the same L2 core + L1 practice with L1 EI (L2+L1 group), or (3) the same L2 core + L1 practice but without L1 EI (L2+L1prac group). Findings indicated that increasing amounts of practice led to more accurate and faster performance only for learners who received L1 EI (L2+L1 group). Coefficient of Variation analyses (Segalowitz & Segalowitz, 1993) indicated knowledge restructuring early on that appeared to lead to gradual automatization over time (Solovyeva and DeKeyser, 2017; Suzuki, 2017). Our findings that EI and practice about L1 processing routines benefited the accuracy, speed, and automaticity of L2 performance have major implications for theories of L2 learning, the role of L1 EI in L2 grammar learning, and L2 pedagogy.

Key words: Explicit instruction, practice, morphosyntax, L1 influence, automaticity, L2 learning

kmcmanus@psu.edu

First language (L1) knowledge and L1 processing routines can heavily influence second language (L2) online processing (Ellis, 2006; Ellis, Hafeeez, Martin, Chen, Boland, & Sagarra 2014; Hopp & Lemmerth, 2016; Roberts & Liszka, 2013) and offline interpretation and production (Ellis & Sagarra, 2011; Huensch & Tracy-Ventura, 2017; Murakami & Alexopoulou, 2016). Several theories of L2 input processing additionally foreground a critical role for L1, such as L1-entrenched attention allocation and blocking (Ellis, 2006; Cintrón-Valentin & Ellis, 2016) and L2 processing routines that can be influenced by the L1 (MacWhinney, 2005, 2012; O'Grady, 2013; VanPatten, 2002). Very little research, however, has examined the extent to which this research evidence base about L1 influence in L2 acquisition can be used to enhance the effectiveness of L2 grammar learning, including theorizing about how explicit information (EI) about the L1 might influence L2 performance, online or offline. Research to date in this area has shown that EI about L1 and L2 form-meaning mappings for crosslinguistically different target features immediately benefitted written, untimed L2 production (Ammar, Lightbown, & Spada, 2010; Horst, White, & Bell, 2010; Kupferberg, 1999), whereas EI about the L2 only (but not about the L1) for crosslinguistically different features did not benefit performance on grammaticality judgment tests (Tolentino & Tokowicz, 2014).

Building on this agenda, McManus and Marsden (2017, 2018) provided EI about the L1 (unlike Tolentino & Tokowicz 2014) and interpretation practice of both French (L2) and English (L1) sentences (unlike any of the aforementioned studies) to investigate their instructional effectiveness for aspect in L2 French, a well-documented area of difficulty due to crosslinguistic differences (Howard, 2005; Izquierdo & Collins, 2008, McManus 2013, 2015). McManus and Marsden's explicit instruction lasted 3.5 hours and was delivered over four weeks. EI about L1 and L2 processing routines followed by interpretation practice of English (L1) and French (L2) sentences improved learners' speed (online) and accuracy (offline) of aspectual interpretation (*Imparfait, Passé Composé, Présent*) four days after instruction (Immediate Posttest) and six weeks later (Delayed Posttest). Whilst that post-instruction evidence suggested that L1 EI benefited L2 online and offline performance, we understand very little about the nature of the actual learning trajectory *during* the practice, including the extent to which learning during the practice was affected by receiving pre-practice EI about the L1. The current study addressed this gap by examining learners' item-by-item interpretation of French sentences while undertaking practice, to better understand how performance during the practice contributed to the learning gains at Immediate Posttest and Delayed Posttest as previously reported by McManus and Marsden (2017, 2018). To our knowledge, no previous research has investigated the extent to which EI about L1 and L2 processing routines can affect the accuracy, speed, and automaticity of learners' responses during practice.

The current study additionally addressed a potential methodological limitation of McManus and Marsden's (2017, 2018) study, in which the two crosslinguistic outcome tests (where items provided an L1 context followed by L2 stimulus) may have advantaged the L2+L1 group. The current study removes this possible confound by examining performance during L2 practice in which *no* L1 context was given in the practice sentences. Thus, benefits for L1 explicit instruction on activities that did not coerce crosslinguistic processing would suggest that McManus and Marsden's previous findings were unlikely to have been an artefact of the nature of the tests themselves.

PRACTICE, AUTOMATIZATION, AND ITS SIGNATURES

Research examining the effectiveness of EI and practice have mostly assessed learning using offline outcome measures, often immediately after instruction without Delayed Posttesting (for review, see Shintani, 2015), with very few analyses of performance during practice. These lines of research cannot (and have not sought to) address theoretical questions about learning sub-processes during practice. Skill Acquisition Theory (Anderson, 1983) proposes sequenced sub-processes that assign practice a key role in development (see also DeKeyser, 2017). First, establishing reliable and accurate declarative knowledge is argued to be essential (Cornillie, Van Den Noorgate, Van den Branden, & Desmet, 2017; DeKeyser, 1997), although no research to date has examined whether providing information about the L1 may affect subsequent stages of skill acquisition. Procedural knowledge is thought to underpin the conscious rule-governed behaviour that rehearses this declarative knowledge and has been characterized by decreasing error rates and faster reaction times. Over time, such practice can lead to automatization, "a fast, parallel, fairly effortless process that is not limited by short-term memory capacity, is not under direct subject control, and is responsible for the performance of well-developed skilled behaviors" (Schneider, Dumais, & Shiffrin 1984, p.1). Although the accuracy and reliability of declarative knowledge representations prior to practice are argued to play a key role (Anderson, 1983), little research exists into longitudinal behavioural signatures that may follow this new declarative knowledge about language, i.e., during practice. Such data are critical for determining the validity of skill acquisition theory in accounting for aspects of L2 learning.

To our knowledge, only two studies have examined fine-grained signatures of learning in longitudinal designs. Both studies found that accuracy improved and reaction times (RTs) decreased as a function of practice. These curves of development showed large changes early on in the practice with smaller changes later on, indicative of qualitative improvements in the stability and efficiency of processing rather than processing speed-up only. In DeKeyser (1997), all participants received the *same* EI about morphosyntax of a novel language and were assigned to one of three practice conditions: comprehension, written production, equal proportions of both. Practice lasted eight weeks, distributed over fifteen sessions (twenty-four practice items per session, with feedback for incorrect responses). Longitudinal analyses across all practice sessions showed that performance was strongly influenced by practice type: "performance in comprehension or production is severely reduced if only the opposite skill was practiced" (p. 213, see also Li & DeKeyser 2017). Furthermore, independent of practice type, DeKeyser found that RTs decreased and accuracy improved as a function of the practice, most noticeably between the first two sessions, with smaller changes between latter sessions. Similar findings were reported by Cornillie et al., (2017), who documented signatures of learning English morphosyntax during online gaming. All participants received the *same* pre-practice EI about the L2, across all treatment groups (as in DeKeyser, 1997), completed the same comprehension practice, but received different types of corrective feedback during the practice: correct/incorrect feedback or correct/incorrect feedback with EI about the L2. Practice was game-based grammaticality judgements over 31 sessions (192 practice items per session) in two practice sessions (with two weeks between them), with two short reading comprehensions before and after gaming. Two target features were investigated: English quantifiers and dative alternation. Results showed similar accuracy scores for both target features in the first practice session. In the second practice session, however, quantifier accuracy scores were higher than those for the dative alternation. In terms of feedback type, additional EI appeared to provide few benefits for dative alternation. Like DeKeyser (1997), within-group analyses showed that increasing amounts

kmcmanus@psu.edu

of practice led to faster and more accurate performance. The largest improvements were also found in the earlier practice sessions, with fewer improvements later on. While both studies considered RT decreases and accuracy increases to reflect processing improvement, the stability of the RT curves of development were considered evidence of automatization leading the authors to conclude that practice led to qualitative improvements in the stability and efficiency of processing behaviour rather than leading only to faster processing.

Because faster RTs could index both automatization (a mechanism within skill acquisition) and, more simply, 'speed-up' (Segalowitz, 2010; Segalowitz & Segalowtiz, 1993), more accurate and faster performance do not necessarily reflect automatic/unconscious processing. Automaticity is the restructuring of underlying processing routines that enhances processing efficiency and stability, but speed-up corresponds to accelerated performance without necessarily indicating qualitative restructuring (Paradis, 2009; Segalowitz, 2010). Segalowitz and Segalowitz (1993) proposed that processing *stability* combined with faster performance may be signatures of greater processing efficiency. To tease apart automatization from processing that speeds up but in the absence of change in the nature of the knowledge, as would be required for proceduralisation and automatization, researchers have used the Coefficient of Variation (CV), a measure of processing stability (mean standard deviation (SD) divided by mean RT). CV distinguishes between a general speed-up (where SDs and RTs decrease at the same rate) and automatization (where the rate of decrease in SDs exceeds the rate of decrease in RTs). This is because automatization is understood to entail elimination or reduction of inefficient sub-processes/components that are the cause of processing variability. Thus, processing *stability* is reflected by SDs of RTs getting narrower over time at a faster rate than the decrease in RTs over time, resulting in a trajectory of decreasing CVs.

CV interpretation in L2 research is mixed. Cross-sectional designs have shown CV reductions as instruction/proficiency level increases (Hulstijn, van Gelderen, & Schoonen, 2009; Lim & Godfroid, 2015), but longitudinal designs have shown that CV trajectories can be more variable with no clear direction of change (Solovyeva & DeKeyser, 2017; Suzuki, 2017). Time is one potential explanation for these findings: longitudinal analyses examined change over hours and days, whereas cross-sectional designs examined change over years. The latter offers more opportunities for practice, understood to be a key driver for automatization, whereas shorter-term yet longitudinal (within-subject) data may reflect earlier stages in skill acquisition: knowledge creation and/or restructuring, as in proceduralization.

To our knowledge, no previous research has used CV signatures following different types of pre-practice EI about morphosyntax to interpret the effects of L2 instruction during practice. One advantage of this design is that we can explore the extent to which CV variability might index creation and/or restructuring of knowledge that is indicative of proceduralization, as suggested by Solovyeva and DeKeyser (2017). Solovyeva and DeKeyser's proposal, however, is based on evidence about lexical processing in a novel language (using lexical decision data of novel words and reanalysis of similar data from Brown & Gaskell 2014 and Bartolotti & Marian 2014). Lim and Godfroid (2015) suggest that CV might better explain lexical processing efficiency because lexical processing tends to rely more heavily on lower level processes (e.g., lexical access), whereas sentence-level/(morpho)syntactic processing tends to require higher level and multi-layered processes, including, for example, lexical access, inferencing, using background information, building a text model (see also Grabe & Stoller 2013). It is possible that CV changes might be more detectable when processing involves fewer component processes (as in lexical processing, for example). Although it has been argued that CV changes (with no clear

kmcmanus@psu.edu

trajectory) might represent signatures of change in the nature of lexical knowledge, the extent to which such CV changes might explain morphosyntactic processing remains an empirical question.

In sum, the current study addresses the following gaps: First, unlike both Cornillie et al. (2017) and DeKeyser (1997) whose learners all received the *same pre-practice* EI about the L2, we compared different types of pre-practice EI: EI about L2 only versus EI about L2 and L1. Second, participants were authentic classroom learners of L2 French, thus contrasting with previous investigations of longitudinal development during practice with (semi-)artificial languages in lab-based settings (but see Cornillie et al.). Third, our instruction focused on the meaning(s) of the grammatical feature under investigation (in contrast to Cornillie et al.). Fourth, we provided extensive practice with many opportunities for proceduralization. Fifth, we examined learners' item-by-item, longitudinal performance during each practice session, thus offering a detailed picture of accuracy and RT trajectories. In these ways, we extend the agenda on using CV as an index of knowledge restructuring and automatization involving morphosyntax.

RESEARCH QUESTIONS

This study examined whether the type of EI (L2-only, L2+L1) provided before practice moderated the accuracy and speed of responses during practice. Faster response speeds, as evidenced by decreasing RTs, were further examined using CV to distinguish between speeded-up and automatic performance. We sought to address the following research questions:

- To what extent do the accuracy, speed, and automaticity (as measured by CV) of responses change over time with increasing amounts of L2 interpretation practice?

- Compared to L2-only EI + interpretation practice, to what extent do the accuracy, speed, and automaticity (as measured by CV) of responses change when undertaking additional L1 interpretation practice with and without L1 EI?

METHOD

Participants

Participants were 53 university learners of French as a foreign language in semester two of a four-year bachelor of arts honours degree in French. All participants were L1 English speakers, aged 18-21, had completed A2-level French (English high school leaving qualification, equivalent to 700 to 800 hours of instruction), and had not spent more than six weeks abroad in a French-speaking country. Mean years of learning French was 10.3 (SD=2.7) and the mean time spent abroad in a French-speaking country was 3.3 weeks (SD=6.07). Advanced-level learners were recruited because our target feature, French *Imparfait*, is acquired late, typically not taught in beginning language classes, and is absent among beginners (Bartning & Schlyter, 2004). Furthermore, in order examine the extent to which different types of EI plus practice can improve learners' knowledge of *Imparfait*'s form-meaning mappings, our design required previous knowledge of IMP's inflectional forms, but not its full set of form-meaning mappings (as was confirmed by Pretest performance[1]).

Target feature: French Imparfait

The target feature was French *Imparfait* (IMP) verbal morphology, a past tense form used to express past habituality and ongoingness (e.g., *il jouait au foot* - 'he used to play/was playing football'). This feature was selected because SLA research has repeatedly shown its full set of functions are late-acquired due to functional complexity, including complex L1-L2 form-meaning mapping differences (see Bartning & Schlyter, 2004; Howard, 2005; McManus, 2013, 2015). 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., run in the park), twelve accomplishments (e.g., walk to the shop) and twelve achievements (e.g., arrive home). For stimuli examples, see Appendix B and IRIS (www.iris-database.org).

Study design

Three instructional treatments were implemented: L2 EI + L2 practice (L2-only, hereafter); L2 EI + L2 practice + L1 EI + L1 practice (L2+L1, hereafter); L2 EI + L2 practice + L1practice (L2+L1prac, hereafter). Participants were randomly assigned to one of these treatments, which were administered one-to-one with laptops using E-Prime 2.0 and delivered in four 45-minute sessions over three weeks, totalling 3.5 hours. The first author collected all of the data.

Sessions one and two were delivered in week two, session three in week three, and session four in week four. There were approximately three days between each session and spacing was the same for each treatment group. In addition, spacing between the final treatment session and the Posttest and Delayed Posttests were almost identical across all treatment groups. (See Suzuki 2017 for a discussion of the potential effects of different distributions of practice and of different ratios of inter-practice and practice-test spacing. As our treatment groups experienced similar spacing, we attempted to control for such effects).

Each session had a different instructional focus on morphemic contrasts expressed by IMP: Session one, ongoingness in the past (IMP) vs. present (Present tense); Session two, habituality in the past (IMP) vs. present (Present tense); Session three, past ongoingness (IMP) vs. past habituality (IMP); Session four, past ongoingness (IMP) vs. past habituality (IMP) vs. past perfectivity (Passé Composé). Critically, sessions one and two presented information that was new (i.e., within the experiment), session three combined information that had already been experienced in sessions one and two, and session four included information that had been experienced in all three previous sessions. All materials are available on IRIS.

Instructional treatments

For all three groups (L2-only, L2+L1, L2+L1prac), treatments included an identical core of EI about French *IMP* and practice interpreting it. We first describe this common core, before describing the additional L1 treatments. Table 1 summarizes the different instructional components received by each treatment group.

EI about L2. Pre-practice EI was first provided for approximately five minutes at the start of each session and depicted conceptual-semantic information using a short video, image, or sound file of events. Ongoingness in present versus past was the instructional focus in session one, for example. Ongoingness was depicted using a ten second video of a man eating an apple, in which the apple was never fully eaten. Learners were then asked to think about how they would describe what they just saw in the video (e.g., he is eating an apple). Then the appropriate L2 aural and written forms were presented, and information given about how to interpret their meaning. For example, French verb endings can be used to distinguish between past ongoingness and present ongoingness (e.g., *il jou<u>ait</u>* – Past IMP, *il jou<u>e</u>* – Present tense), and so watching/listening out for verb endings can be helpful to distinguish ongoingness in the past vs. present in French. See Appendix C for description of EI used in session one.

Practice in L2. Pre-practice EI was immediately followed by practice in listening and reading that forced learners to attend to form-meaning mappings expressed by *IMP*, *Passé Composé* or *Présent* (see VanPatten 2002 for referential activities in Processing Instruction, and Marsden 2006 on using inflections to interpret tense). Learners selected the stimulus's meaning from two options in sessions one to three and three options in session four. The L2 practice contained 552 exemplars (96 in each of sessions one and two, 144 in session three, and 216 in session four), that were randomly ordered within each session for different participants; each verb type occurred eight times with *IMP* (n = 384), counterbalanced across listening/reading and ongoing/habitual[2]. All learners completed the same amounts of L2 practice across all treatments, though the items within each practice session were presented randomly by E-Prime. See Appendix A for frequencies of French stimuli and examples. Stimuli were single clause in sessions one and two (e.g., *Il court dans la rue* 'he is running in the street). To practice interpreting IMP's habituality or ongoingness by relying, critically, on the inflectional morphemes in the broader discourse context, two clause stimuli were necessary in sessions three and four (e.g., *Elle mangeait un sandwich quand la cloche a sonné* ('She was eating a sandwich when the bell rang'). An image (e.g., sandwich) plus a bracketed infinitive (e.g., *manger* 'eat') appeared alongside two-clause stimuli so that learners knew which verb to interpret. The stimulus appeared first (e.g., *jouait au foot quand sa petite amie est arrivee* 'was playing football when his girlfriend arrived'), then after 2500ms (for two-clause stimuli) and 500ms (for single-clause stimuli) the response options appeared and stayed on screen until a response was pressed.

For aural stimuli, the response options did not appear until after the full stimuli had played. Thus, for all practice items, responses were not time pressured. Responses could not be changed after initial selection.

Correct/incorrect feedback was shown immediately after each response. Additional EI was provided during the practice following incorrect responses only, which, as Appendix A, Table A2, shows, was infrequent and occurred in very (statistically) similar amounts in all treatments.

L2+L1 treatment. In addition to EI about L2 and practice in L2, the L2+L1 treatment included *pre-practice EI about the L1* (how English expresses the meanings taught in each session, e.g. ongoingness in session one) as well as practice interpreting L1 forms expressing those same meanings (e.g., present versus past progressive in session one). The design of the L1 EI and L1 practice followed the exact same design principles as outlined above for L2 EI and L2 practice. See Appendix C for description of L1 EI used in session one, and Table A3, for frequencies of English stimuli[3].

Correct/incorrect feedback was shown after each response. Additional EI was given following incorrect responses only[4].

L2+L1prac treatment**.** The L2+L1prac treatment included L2 EI and L2 practice (as in L2-only and L2+L1 treatments) plus L1 practice. No EI about the L1 was provided, either pre-practice or during the practice.

[Table 1 here]

Data analysis

E-Prime collected accuracy and RT data for every response. For accuracy, responses were coded as correct (1) or incorrect (0). Reliability coefficients for accuracy, calculated using the Kuder-Richardson Formula 20, were: Session one (.91), Session two (.87), Session three (.73), and Session four (.79)[5]. RTs were calculated in milliseconds from the onset of response options to response selection. We analyzed raw RT data, trimmed in line with Keating & Jegerski's (2015) recommendations, removing RTs less than 150ms and greater than 2,000ms. RT. Verbs were coded verbs as irregular or regular.

Accuracy and RT were analysed separately in R (R Core Team 2018). Also, separate analyses were conducted for each session because each had a different instructional focus (as previously described) and not all sessions included the same number of practice items. For accuracy, we conducted logit mixed-effects analyses (Jaeger, 2008) using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015). For RT, we conducted mixed-effects linear regression analyses (Baayen, Davidson, & Bates 2008) using *nlme* (Pinheiro, Bate, DebRoy, Sarkar, & R Core Team, 2018). For both accuracy and RT analyses, explanatory variables were as follows: Group (L2+L1, coded as -1; L2+L1prac, coded as 1; L2-only, coded as 0); Item (i.e. ranked practice item number); and Verb (regular, coded as 1; irregular, coded as 0). These were entered into the models as fixed effects. Subject and items were added as cross-random factors.

In contrast to ANOVAs, mixed-effects analyses avoid violating the assumption of independence because they model relationships between observations, an important consideration for our longitudinal analyses (Field, Miles, & Field 2012; Murakami 2016). Mixed-effects models additionally offer many other advantages over ANOVAs, including greater flexibility of data distribution (e.g., binomial variables) and robustness against violations of homoscedasticity and sphericity. This makes mixed-effects models particularly useful for longitudinal research and more desirable for our analyses (Cunnings & Finlayson 2015; Linck & Cunnings 2015).

For each session, multiple models were constructed and the most plausible model was found through comparison. We started with the simplest model, with new parameters added to the model one at time (Field et al. 2012; Murakami 2016). We compared models as they were built using maximum-likelihood estimation (Field et al. 2012).

First, we fitted a base-line model in which we included only the intercept, then we fitted a model that allowed the intercept to vary over Subjects. Finally, to verify whether allowing the intercepts to vary improved the model fit significantly, we compared the models using AIC (Akaike Information Criterion) and the *anova* function. The final models were then built by adding Group, Item, and Verb as fixed-effect factors, followed by a random slope added for the effect of Item (thus allowing the effect of Item to vary across Subjects, because items were randomly ordered within each session for each participant), and then a Group x Item fixed-effect interaction. After adding each new parameter to the model, we verified whether its addition significantly improved the fit of the model (using AIC and *anova*, see Table 2 for accuracy and Table 3 for RT). A parameter was only retained in the optimal model if its addition significantly decreased the AIC value (see Cunnings & Finlayson, 2015; Field et al. 2012). For example, the Verb parameter in Sessions 1-3 for RT did not significantly improve the final model, but its removal did. As a result, our optimal models in Sessions 1-3 for RT excluded the Verb parameter.

Because each optimal model contained three treatment groups, Group x Item interactions were further explored using *lme4* (for accuracy) and *nlme* (for RT) for each Group (equivalent to posthoc testing, see Field et al. 2012), thus allowing further examination of treatment effects on performance over time.

For RTs that significantly quickened over time in each session, we calculated the Coefficient of Variation (CV, mean SD divided by mean RT). As done by Hulstijn et al. (2009), Lim and Godfroid (2015), and Suzuki and Sunada (2018), data for our CV analyses included RTs for correct responses only, and excluded (a) incorrect responses, to reduce potential confounds between processing speed and accuracy of linguistic knowledge, and (b) extremely slow RTs of more than three SDs above the mean, to exclude potentially invalid data. This procedure removed 3272 data points (11.3% of the data). Simple linear regression analyses were used to model the nature and size of the relationship between CV (outcome variable) and ranked item number (predictor variable). Linearity was examined using scatterplots, which showed linear distribution of the data.

For all analyses, the alpha was set at .05. To interpret effect estimates and magnitudes of change, we present 95% CIs and R^2 effect sizes. CIs that do not pass through zero can be considered reliable indicators of change. Like other standardized effect size statistics, R^2 can be used as a summary index for statistical models to evaluate model fit, compare magnitudes of effect across studies, and can be used for meta-analysis (Nakagawa and Schielzeth, 2013). R^2 values range from 0-1 and are used to estimate how much of the variance in performance (accuracy, RT, and CV) can be accounted for by Group, Item (ranked item number), and Verb (ir/regularity), individually and collectively (see Plonsky & Oswald 2017). We report R^2 values for all fixed effects (marginal R^2), computed using the *MuMIn* package (Bartoñ 2018) in R (R Core Team, 2018). R^2 values around .18, .32, and .51 are interpreted as small, medium, and large, respectively, in terms of the explained variance they represent (Plonsky & Ghanbar, 2018).

kmcmanus@psu.edu

[Table 2]

[Table 3]

RESULTS

Results are presented separately for accuracy and RT. CV analyses are used to interpret RTs that reduced over time.

Accuracy

Table 4 shows the effects of the fixed factors and the interaction between treatment group and ranked item number for accuracy in all practice sessions (see also Figure 1 for corresponding plots with 95% CI shading).

Verb regularity did not significantly influence the accuracy of learners' performance in any practice session (p >. 05, all CIs passed through zero). Group, item number, and the interaction between group and item number, however, were all statistically significant (with CIs that did not pass through zero), indicating that group and item number both individually and together significantly influenced the accuracy of learners' performance over time.

Posthocs tests examining each group's performance over time indicated that the L2+L1group's accuracy improved significantly over time in all four practice sessions (session one, b = .07 [95% CI: .03, .09], $z = 4.54$, $p < .001$, R^2 = .03; session two, b = .05 [95% CI: .02, .08], $z = 3.24$, p = .001, R^2 = .04; session three, b = .04 [95% CI: .03, .05], $z = 5.85$, $p < .001$, R^2 = .04; session four, b = .02 [95% CI: .01, .02], $z = 7.48$, $p < .001$, R^2 = .08). Although R^2 values over the four practice sessions were very small overall, practice explained more of the variance in performance in session four than any of the previous sessions.

In contrast, we found that accuracy did not significantly improve over time for the L2+L1prac and L2-only groups. For L2+L1prac, accuracy worsened slightly but significantly over time in Session four (*b* = -.00 [95% CI: -.01, -.00], *z = -2.32, p* = .02, R^2 = .01), but not in the other practice sessions (session one, *b* = -.01 [95% CI: -.01, .03], *z = -1.68, p* = .09, R^2 = .00; session two, *b* = -.01 [95% CI: -.03, .01], *z = -1.15, p* = .25, R^2 = .01; session three, *b* = -.00 [95% CI: -.00, .01], *z = .77, p* = .44, R^2 = .01). Results for the L2-only group showed no change over time for accuracy in any of the sessions (session one, *b* = .01 [95% CI: -.01, .04], *z = 1.09, p* = .27, R^2 = .01; session two, *b* = -.01 [95% CI: -.03, .02], *z* = -*.54, p* = .59, R^2 = .01; session three, *b* = .00 [95% CI: -.01, .01], *z = .82, p* = .41, R^2 = .01; session four, *b* = .00 [95% CI: -.01, .00], *z = .06, p* = .95, R^2 = .02). All CIs for L2+L1prac and L2-only either passed through zero and/or included zero. There were also few changes in R^2 values over time, which were even smaller than those found for the L2+L1 treatment, indicating that increasing amounts of practice contributed little to explaining performance, thus contrasting with the patterning of results found for the L2+L1 treatment.

Taken together, these results indicate that only the L2+L1 group's accuracy over time significantly improved with increasing amounts of practice. We found no such evidence for the L2+L1prac and L2-only groups. These learning trajectories are visualized in Figure 1.

[Table 4 here]

[Figure 1 here]

Reaction times

Table 5 shows the effects of the fixed factors and the interaction between treatment group and ranked item number for RT in all practice sessions (see also Figure 2 for corresponding plots with 95% CI shading).

The addition of a fixed main effect for Verb in sessions one, two, and three did not lead to an improvement of model fit (see Table 3), indicating that verb regularity did not significantly influence the speed of learners' performance in these sessions. In session four, however, verb regularity significantly influenced the speed of learners' performance (p <. 05, CIs did not pass through zero). Although posthoc tests showed significantly slower performance on irregular than regular verbs for L2+L1prac (*b* = -113.81 [95% CI: -184.62, -43.00], *t*(4083) = -3.15, *p* = .002, R^2 = .09) and L2-only (*b* = -75.97 [95% CI: -151.53, -.41], *t*(3653) = -1.97, *p* = .05, R^2 = .04), verb regularity only explained a very small proportion of the variance. In the L2+L1 group, however, we found no differences between irregular and regular verbs (*b* = -14.89 [95% CI: -75.06, 45.28], *t*(3653) = -.49, *p* = .63, R^2 = .04).

Group, item number, and the interaction between group and item number, however, were all statistically significant (with CIs that did not pass through zero), indicating that group and item number both individually and together significantly influenced the speed of learners' performance over time.

Posthocs tests examining each group's performance over time indicated that the L2+L1group's speed of performance got significantly faster over time in all four practice sessions (Session 1, *b* = -10.27 [95% CI: -12.12, -8.27], *t*(1614) = -10.96, *p* < .001, R^2 = .13; Session 2, *b* = -9.61 [95% CI: -12.09, -7.13], *t*(1614) = -7.58, *p* < .001, R^2 = .12; Session 3, *b* = -8.37 [95% CI: -9.56, -7.18], *t*(2430) = -13.74, *p* < .001, R^2 = .20; Session 4, *b* = -6.59 [95% CI: -

7,73 -5.47], $t(3653) = -11.48$, $p < .001$, $R^2 = .32$). R^2 values over the four practice sessions additionally indicated that practice explained more of the variance in performance in session four than any of the previous sessions, similar to our findings for accuracy, albeit with larger R^2 values (e.g. session four R^2 values were .08 for accuracy, but .32 for RT).

In contrast, we found that L2+L1prac's tended not to change significantly over time, except in session three when RTs got significantly faster over time ($b = -1.92$ [95% CI: -3.05, -.80], $t(2716) = -3.36$, $p < .001$, $R^2 = .02$), but we found no significant change in the other sessions (Session 1, $b = -.44$ [95%CI: -2.61, 1.73], $t(1804) =-.39$, $p = .69$, $R^2 = .06$; Session 2, $b = -.20$ [95% CI: -3.04, 2.63], $t(94) = -.14$, $p = .89$, $R^2 = .01$; Session 4, $b = -.51$ [95% CI: -2.02, .99], $t(4083) = -66$, $p = .51$, $R^2 = .09$). Except in session three, CIs passed through zero. Similarly, L2-only's speed of performance did not change significantly over time (Session 1, $b = .38$ [95%CI: -2.75, 3.50], $t(1614)= .24$, $p = .81$, $R^2 = .09$; Session 2, $b = -.08$ [95% CI: -3.50, 3.34], $t(1614) = -.05$, $p = .96$, $R^2 = .02$; Session 3, $b = -75$ [95% CI: -3.19, 1.70], $t(2287) = -.59$, $p = .55$, $R^2 = .07$; Session 4, $b = -1.23$ [95% CI: -2.62, .15], $t(3653) = -1.75$, $p = .08$, $R^2 = .04$). CIs in all sessions passed through zero, and there was no clear trajectory of R^2 values.

Consistent with our results for accuracy, fixed main effects for group, item number, and the interaction between group and item number were all statistically significant in all four Sessions ($p < .05$, CIs did not pass through zero), suggesting that group and item number, both individually and together, significantly influenced the speed of learners' performance over time. Exclusively in session four, verb regularity significantly influenced L2+L1prac's and L2-only's reaction times, in that they were slower at giving responses to irregular than regular verbs. Verb regularity did not influence L2+L1's performance. In sum, therefore, the L2+L1 group's performance over time significantly improved as a function of the practice. We found no such evidence for the L2-only group. L2+L1prac's RT got faster over time in session three, but there were no changes in sessions one, two and four. Practice explained a medium-sized proportion of the variance in Session 4 for the L2+L1group. These learning trajectories are visualized in Figure 2.

[Table 5 here]

[Figure 2 here]

Automaticity as measured by Coefficient of Variation of reaction times

Because only L2+L1's RTs decreased significantly over time, we present CV analyses to interpret the faster RTs in this treatment group (for summary, see Figure 2). Recall, CVs scores that increase, remain broadly constant, or have no clear direction have traditionally been argued to indicate speed-up or they may indicate, as argued more recently by Solovyeva & DeKeyser (2017) in relation to novel word learning, the creation and/or restructuring of knowledge. CVs that gradually decrease over time are thought to indicate qualitative changes in processing efficiency and stability, indicative of automatization, a process driven by practice.

In order to ascertain the extent to which CVs significantly reduced over time, linear regressions were calculated to predict CVs based on item number (see Table 6). Results showed that CVs in sessions one and two were variable with no clear direction over time: CV trajectories over time were broadly bell shaped (session one) or "S" shaped (session two). Ranked item number was not a significant predictor of CVs in session one (R^2 = .00). In session two, however, ranked item number was a significant predictor of increasing CVs (R^2 = .09).

In sessions three and four, reductions in CV appear more visible (see Figure 3). Linear regression results in these sessions showed that ranked item number significantly predicted

kmcmanus@psu.edu

decreasing CVs (session three, $R^2 = .25$; session four, $R^2 = .32$). In other words, CVs reduced with increasing amounts of practice. Increasing R^2 values indicate that item number explained more of the variance in Session 4 than in Session 3, and in both sessions, practice explained a small-to-medium proportion of the variance.

In sum, CV trajectories in sessions one and two had no clear direction or changed little over time (small R^2 value in session two), and then appeared to visibly and reliably decrease over time in sessions three and four (medium R^2 values). These results suggest that processing efficiency and stability, indicative of automatization, was not evident in the earlier sessions and was only observable in the last two practice sessions.

[Table 6 here]

[Figure 3 here]

DISCUSSION

We examined whether EI about the L1 and/or practice in interpreting the L1 affected the accuracy and speed of learners' responses during L2 practice, in comparison to receiving only instruction about the L2 (EI and practice). All groups received the same L2 instruction. We examined fine-grained item-by-item performance over time for accuracy and RT over the course of four practice sessions.

Results showed that increasing amounts of practice led to more accurate and faster performance in the group that received L1 EI (L2+L1), but not in the groups that did not (L2+L1prac, L2-only).

Since all groups received the same L2 EI and practice, these results indicate that L2 practice alone did not lead to the differences observed. This contrasts with the findings of Cornillie et al. (2017) and DeKeyser (1997), where automatization effects were detected after L2 explicit instruction. However, different study designs may explain the difference in our results compared with those from previous research. First, Cornillie et al. provided larger amounts of corrective feedback (yes/no plus EI). Second, DeKeyser's practice was distributed over a longer period of time (his fifteen weeks vs. our three weeks). Third, the L1 EI and practice may have been necessary in our study. That is, since neither Cornillie et al. nor DeKeyser focused on crosslinguistically complex form-meaning mappings, additional L1 EI may have been necessary for our target feature (*IMP*) due to its crosslinguistic complexity. In particular, it may have been necessary to elicit subtle changes among these upper intermediate-advanced learners who were already relatively accurate, at least in terms of the target form (but not its form-meaning mappings).

Although our plots showing performance over time showed initially lower accuracy and slower RTs for L2+L1 than the other groups in each session, L2+L1's performance significantly improved with practice. These trajectories indicate that L1 EI (received only by the L2+L1 group) created a delayed advantage: performance was initially less accurate and slower but increasing amounts of practice led to more accurate and faster performance than in the groups without L1 EI.

We used CVs to interpret L2+L1's faster RTs over time. CV trajectories had no clear direction during sessions one and two and decreased during sessions three and four. Recall also that sessions one and two presented and practiced information that was new for the participants (within the context of this experiment), and sessions three and four revisited this information through different configurations of practice. Solovyeva and DeKeyser (2017) proposed two interpretations for CV change. First, knowledge creation and/or restructuring is reflected in CV trajectories that vary but with no clear direction because new (sub)processes are added to existing processing routines (see also Suzuki, 2017). Second, automatization of established/existing knowledge results in *decreasing* CVs due to the elimination and/or restructuring of inefficient processing routines (see also Hulstijn et al., 2009; Lim & Godfroid, 2014). Solovyeva and DeKeyser's (2017) proposals repurposed CV, when observed in different patterns, as an indicator both of learning in the *earlier* stages in skill acquisition (where declarative knowledge is established and incorporated into existing knowledge, as it is proceduralized), as well as of the *later* stages of automatization. Although, to our knowledge, no previous research has used CVs to examine the effects of different types of pre-practice EI on L2 performance during practice, Solovyeva and DeKeyser's (2017) hypothesis helps explain our observed trajectories.

First, CV trajectories that vary with no clear direction, as found in sessions one and two, in particular, suggests the restructuring of existing L2 knowledge through the addition of new processes and/or representations (Solovyeva & DeKeyser 2017). In our case, pre-practice EI about the L1 provided new information about form-meaning mappings and processing routines for ongoingness and habituality. We suggest that CV trajectories with no clear direction were underpinned by the integration of new knowledge (EI about the L1) with existing L2 knowledge, thereby changing the nature of the L2 knowledge and its processing, and introducing temporary instabilities into the L2 knowledge system (see Figure 3).

Second, CV decreases in Sessions 3 and 4 appear compatible with automatization of knowledge, due to the elimination of slower, less efficient processing procedures. Our results indicate that reducing CVs only emerged after opportunities to undertake considerable practice (approximately 1.5 hours over two prior sessions, that introduced and rehearsed the same information though in different types of practice items). Indeed, Session 4 contained the most practice items, and rehearsed information that had already been presented and practiced in three prior sessions, which could explain why we see clearer CV decreases because there were more opportunities for automatization to occur, both within the session and prior to it.

These accuracy, RT, and CV results for performance during practice are consistent with McManus and Marsden's (2017, 2018) previously-discussed post-instruction findings at Posttest and Delayed Posttest, which showed that providing L1 EI with L1 practice alongside a core of L2 EI with L2 practice (i.e., the L2+L1 treatment) improved the speed (on a self-paced reading task) and accuracy (on a sentence judgement task in reading and listening) of L2 processing, immediately after instruction and with gains retained six weeks later. There were few reliable learning benefits for groups that did not receive L1 EI as part of their instruction.

Taken together, two trends emerge from the current study's findings and those for post-instruction performance as reported by McManus and Marsden (2017, 2018). First, performance during the practice was consistent with performance at both Posttest and Delayed Posttest: learners whose performance improved during practice also showed improvement in the outcome measures. Second, our findings indicate that improvement in the accuracy, speed, and stability of L2 performance, both during the practice and post-instruction at the posttests, was found only for learners whose treatment included EI about the L1 (i.e., the L2+L1 group). In other words, L2 practice by itself (without EI about the L1, and even if accompanied by practice in the L1), did not improve the accuracy, speed, and stability of L2 performance, either during the practice or post-instruction at the posttests.

Learning processes during practice

Our results suggest that the L2+L1 treatment (additional L1 EI + L1 practice) played an important role in improving the accuracy, speed, and stability of learners' responses during practice. L2+L1 EI was more effective than L2-only EI arguably because it addressed the nature of the crosslinguistic learning problem.

Our CV results indicated qualitative changes in learners' processing, suggesting reduction/elimination of inefficient sub-processes/components that are understood to be a cause of processing variability (Segalowitz, 2010). Over time, systematic practice appeared to lead to more efficient/stable processing, in line with our pedagogical aims. We speculate that CV trajectories that vary with no clear direction in sessions one and two followed by CV reductions over time in sessions three and four reflected adjustment of prior processing routines. We further speculate that this constituted moving away from routines that interpreted IMP via L1 processing routines, such as via a constrained mapping of IMP to meanings expressed by L1 forms, like 'BE(past) + ing' (for ongoing meaning) or 'used to + verb' (for habitual meaning) or via lexical cues (e.g., adverbials), towards routines that more speedily and reliably interpreted IMP by using inflectional morphology that is used elsewhere in the sentence as a reliable cue for extracting the habitual and ongoing meanings of IMP. This would be consistent with some interpretations that decreases in CV indicate greater processing stability and efficiency brought about by extensive opportunities for practice (e.g., Solovyeva & DeKeyser, 2017; Suzuki, 2017).

LIMITATIONS AND FUTURE RESEARCH

Given our small sample sizes, we emphasize that our accounts are tentative. Also, our interpretations that CV trajectories that vary with no clear direction index knowledge creation and/or restructuring rest on a small body of evidence and more research is needed to corroborate these interpretations. Nonetheless, this constitutes an important research agenda if we wish to understand the mechanisms underpinning learning effects during practice and seek empirical support for skill acquisition accounts of learning L2s.

We leave to future research the task of investigating how these signatures of automaticity relate to comprehension and production performance *after* instruction. According to the post-practice performance results in McManus and Marsden (2017, 2018), it seems that performance on a controlled, interpretation outcome measure (self-paced reading test), even six weeks post-instruction, was in line with the during-practice trends observed in the current analysis. That is, both the post- and during-instruction measures showed the most benefits for the group that received additional L1 EI plus practice.

The current study provides evidence of benefits of L1 EI (combined with L1 practice, L2 EI, and L2 practice) on L2 inflectional verb morphology with, specifically, crosslinguistically different form-meaning mappings. We saw that a CV signature of automatization was observable most clearly and reliably during a fourth training session (after 2.25 hours of training). Perhaps critically, this fourth session facilitated repetitive interpretation practice of the material introduced in the previous three sessions. During those first three sessions, and particularly the first two, we observed evidence indicative of knowledge creation and/or restructuring. It remains to be determined how much practice is required for evidence of automatization to emerge, beyond a general speeding-up, for other features and L2 proficiencies.

Our finding that additional L1 EI benefitted the learning of a crosslinguistically complex L2 feature provides some evidence of the usefulness of explicit L1 grammar teaching. For this target feature, L1 grammar teaching led to evidence of change in the nature of L2 knowledge and the speed of access to it. Future research should investigate the extent to which L1 EI may benefit the L2 learning of other linguistic features. Based on these data and our previous research (see McManus & Marsden 2017, 2018), however, L1 EI might be most beneficial for L2 features that are sensitive to crosslinguistic influence at the level of form-meaning mapping (as determined by SLA research), especially for target-features that exhibit L1-L2 form-meaning mapping differences like IMP for English-speakers. Other target features could include learning of (1) the *ser*-*estar* distinction in L2 Spanish and (2) *zai* in L2 Chinese by L1 English speakers. Similar to IMP, *ser*-*estar* and *zai* exhibit complex L1-L2 form-meaning mapping differences: (1) the meaning expressed by a single form in English (*be*) is expressed by multiple forms in Spanish (*ser* and *estar*; see Silva-Corvalán 2014), and the distinction between 'permanent' and 'temporary' characteristics (which *ser* and *estar* convey) can be expressed by multiple adjectival forms in English (*bored* and *boring*) and one form in Spanish (*aburrido*) or (2) the meaning expressed by a single form in Chinese (*zai*) is expressed by multiple forms in English (*progressive V+ing* and *prepositional 'in'*, see Xiao & McEnery 2004). Future research should also investigate the extent to which the usefulness of L1 EI for L2 learning is moderated by age and L2 proficiency. These avenues would help tailor L2 instruction to the nature of the learning problem in different contexts.

REFERENCES

Ammar, A., Lightbown, P., & Spada, N. (2010). Awareness of L1/L2 differences: Does it matter? *Language Awareness, 19,* 129-146.

Anderson, J. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22,* 261-295.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language, 59*, 390-412.

Bartning, I., & Schlyter, S. (2004). Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies*, *14*, 281-299.

Bartoñ, K. (2018). *MuMIn: Multi-Model Inference*. R package version 1.40.4. https://CRAN.R-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67,* 1-48.

Bartolotti, J. & Marian, V. (2014). Wordlikeness and novel word learning. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 146-151). Austin, TX: Cognitive Science Society.

Brown, H., & Gaskell, G. (2014). The time-course of talker-specificity and lexical competition during word learning. *Language, Cognition and Neuroscience, 29,* 1163-1179.

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.

Cornillie, F., Van Den Noortgate, W., Van den Branden, K., & Desmet, P. (2017). Examining focused L2 practice: From in vitro to in vivo. *Language Learning & Technology*, *21*, 121-145.

Cunnings, I., & Finlayson, I. (2015). Mixed effects modeling and longitudinal data analysis. In Plonsky, L. (ed). *Advancing Quantitative Methods in Second Language Research* (pp. 159-181). New York: Routledge.

DeKeyser, R. (1997). Beyond explicit rule learning. *Studies in Second Language Acquisition*, *19,* 195-221.

DeKeyser, R. (2015). 'Skill acquisition theory'. In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 94-112). New York: Routledge.

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., & 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., 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.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics using R*. London: Sage Publications.

Grabe, W. P., & Stoller, F. L. (2013). *Teaching and Researching: Reading*. New York: Routledge.

Hopp, H., & Lemmerth, N. (2016). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition,* 1-29.

Howard, M. (2005). Les contextes prototypiques et marquées de l'emploi de l'imparfait par l'apprenant du français langue étrangère. In E. Labeau & P. Larrivée (Eds.), *Nouveaux développements de l'imparfait*. Amsterdam : Rodopi.

Horst, M., White, J., & Bell, P. (2010). First and second language knowledge in the language classroom. *International Journal of Bilingualism, 14,* 331-349.

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

Hulstijn, J. H., Van Gelderen, A., & Schoonen, R. (2009). Automatization in second language acquisition: What does the coefficient of variation tell us?. *Applied Psycholinguistics, 30,* 555-582.

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.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language, 59*, 434-446.

Keating, G., & Jegerski, J. (2015). Experimental designs in sentence processing research. *Studies in Second Language Acquisition, 37,* 1-32.

Kupferberg, I. (1999). The cognitive turn of contrastive analysis: Empirical evidence. *Language Awareness, 8*, 210-213.

Lim, H. & Godfroid, A. (2015). Automatization in second language sentence processing: A partial, conceptual replication of Hulstijn, Van Gelderen, and Schoonen's 2009 study. *Applied Psycholinguistics, 36,* 1247-1282.

Linck, J. A., & Cunnings, I. (2015). The Utility and Application of Mixed-Effects Models in Second Language Research. *Language Learning*, *65*, 185-207.

MacWhinney, B. (2005). A unified model of language acquisition. In J.F. Kroll & A.M.B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. *49-67*). Oxford: Oxford University Press.

MacWhinney, B. (2012). The logic of the unified model. In S.M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 211-227). New York: Routledge.

Marsden, E. (2006). Exploring input processing in the classroom: An experimental comparison of processing instruction and enriched input. *Language Learning*, *56*, 507-566.

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: A comparison of English and German learners of French. *Canadian Modern Language Review, 71,* 51-77.

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.

Murakami, A. (2016). Modeling systematicity and individuality in nonlinear second language development: The case of English grammatical morphemes. *Language Learning, 66*, 834-871.

Murakami, A., & Alexopoulou, T. (2016). L1 influence on the acquisition order of English grammatical morphemes. *Studies in Second Language Acquisition, 38,* 365-401.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133-142.

O'Grady, W. (2013). The illusion of language acquisition. *Approaches to Bilingualism*, *3*, 253-85.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R Core Team (2018). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-131.1. https://CRAN.R-project.org/package=nlme.

Shintani, N. (2015). The effectiveness of processing instruction and production-based instruction on L2 grammar acquisition: A meta-analysis. *Applied Linguistics*, *36*, 306-325.

Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages*. John Benjamins.

Suzuki, Y. (2017). The role of procedural learning ability in automatization of L2 morphology under different learning schedules: An exploratory study. *Studies in Second Language Acquisition*, 1-15.

Plonsky, L., & Derrick, D. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal*, *100*, 1-15.

Plonsky, L, & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R^2 values. *The Modern Language Journal*.

Plonsky, L., & Oswald, F. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878-912.

Plonsky, L., & Oswald, F. L. (2017). Multiple regression as a flexible alternative to ANOVA in L2 research. *Studies in Second Language Acquisition*, *39*, 579-592.

R Core Team (2018). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna, Austria. https://www.R-project.org/.

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.

Schneider, W., Dumais, D., & Schiffrin, R. (1984). Automatic and control processing and attention. In R. Parasuraman (ed.): *Varieties of attention*. New York: Academic Press.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.

Segalowitz, N., & Segalowitz, S. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics, 14,* 369-385.

Solovyeva, K., & DeKeyser, R. (2017). Response time variability signatures of novel word learning. *Studies in Second Language Acquisition,* 1-15.

Suzuki, Y. (2017). The role of procedural learning ability in automatization of L2 morphology under different learning schedules. An exploratory study. *Studies in Second Language Acquisition* 1-15.

Suzuki, Y., & Sunada, M. (2016). Automatization in second language sentence processing: Relationship between elicited imitation and maze tasks. *Bilingualism: Language and Cognition, 21*, 32-46.

Tolentino, L. & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning, 64,* 279-309.

VanPatten, B. (2002). Processing instruction: An update. *Language learning*, *52*, 755-803.

Xiao, R., & McEnery, T. (2004). *Aspect in Mandarin Chinese: A corpus-based study*. Amsterdam: John Benjamins.

Notes

1. Baseline parity using Pretest scores was tested from the context-matching tests in listening and reading and a self-paced reading test. Baseline parity was found between all groups on all measures. For full reporting of these results, see McManus and Marsden (2017, 2018).
2. Aural stimuli were recorded by two native French speakers. The French sentences were verified for authenticity by 26 native French speakers: All were rated as 100% acceptable, with the meanings (ongoing/ habitual, present/past) as intended by the researchers.
3. Performance on L1 practice items is not analysed here as we were interested in the effects of different types of pre-practice EI on *L2* learning during practice.
4. EI provided during practice provided information about L1 and L2 form-meaning mappings.
5. Based on a meta-analysis of reliability coefficients in L2 research, Plonsky and Derrick (2016) propose that .74 should be considered a general (not absolute) threshold for an acceptable estimate of instrument reliability.
6. Individual data points are not plotted due to the binary nature (0, 1) of the data coding

TABLES

Table 1.
Summary of instructional differences between the treatment groups

	L2+L1	L2+L1prac	L2-only
L2 practice	✓	✓	✓
L2 EI	✓	✓	✓
L1 practice	✓	✓	
L1 EI	✓		

Table 2. Summary of logit mixed-effects model comparisons for accuracy

Session #	Model	Fixed Effects	Random Effects	AIC	Δ AIC	-2LL Statistic	*p*
1	Model 1	None	By-Subject random-intercepts	818.87			
	Model 2	Model 1 + Group	Same as Model 1	820.18	-407.09	$X^2(1) = .69$	.41
	Model 3	Model 2 + Item	Same as Model 1	796.91	-394.45	$X^2(1) = 25.27$	< .001
	Model 4	Same as Model 3	Model 1 + by-Item random slope	798.84	-393.42	$X^2(2) = 2.07$	.36
	Model 5	Same as Model 4 + Verb	Same as Model 4	800.12	-393.06	$X^2(1) = .72$	.39
	Model 6	Model 5 + Group x Item interaction	Same as Model 4	792.58	-388.29	$X^2(1) = 9.53$	.002
	Model 7	Model 4 + Group x Item interaction	Same as Model 4	791.28	-388.64	$X^2(1) = .69$	.40
2	Model 1	None	By-Subject random-intercepts	940.99	0468.50		
	Model 2	Model 1 + Group	Same as Model 1	942.58	-468.29	$X^2(1) = .41$	.52
	Model 3	Model 2 + Item	Same as Model 1	939.82	-465.91	$X^2(1) = 4.76$	.03
	Model 4	Same as Model 3	Model 1 + by-Item random slope	936.01	-462.01	$X^2(1) = 7.81$	.02
	Model 5	Same as Model 4 + Verb	Same as Model 4	937.60	-461.80	$X^2(1) = .41$	.52
	Model 6	Model 5 + Group x Item interaction	Same as Model 4	927.09	-455.55	$X^2(1) = 12.51$	<. 001
	Model 7	Model 4 + Group x Item interaction	Same as Model 4	925.49	-455.49	$X^2(1) = .39$	.53
3	Model 1	None	By-Subject random-intercepts	1776.6	-886.30		
	Model 2	Model 1 + Group	Same as Model 1	1775.0	-884.51	$X^2(1) = 3.58$	.06
	Model 3	Model 2 + Item	Same as Model 1	1749.4	-860.70	$X^2(1) = 27.63$	<. 001
	Model 4	Same as Model 3	Model 1 + by-Item random slope	1730.8	-859.38	$X^2(2) = 22.63$	< .001
	Model 5	Same as Model 4 + Verb	Same as Model 4	1731.8	-858.91	$X^2(1) = .95$	.33
	Model 6	Model 5 + Group x Item interaction	Same as Model 4	1717.1	-850.55	$X^2(1) = 16.71$	< .001
	Model 7	Model 4 + Group x Item interaction	Same as Model 4	1716.1	-851.04	$X^2(1) = .97$	.32
4	Model 1	None	By-Subject random-intercepts	5766.6	-2881.3		
	Model 2	Model 1 + Group	Same as Model 1	5767.9	-2881.0	$X^2(1) = .72$	.39
	Model 3	Model 2 + Item	Same as Model 1	5734.1	-2863.1	$X^2(1) = 35.80$	< .001
	Model 4	Same as Model 3	Model 1 + by-Item random slope	5623.1	-2805.6	$X^2(2) = 114.99$	< .001
	Model 5	Same as Model 4 + Verb	Same as Model 4	5624.5	-2805.2	$X^2(1) = .64$	.42
	Model 6	Model 5 + Group x Item interaction	Same as Model 4	5583.2	-2783.6	$X^2(1) = 43.31$	< .001
	Model 7	Model 4 + Group x Item interaction	Same as Model 4	5581.8	-2783.9	$X^2(1) = .66$	.42

Note. Grey shading indicates optimal model

Table 3.
Summary of mixed-effects linear model comparisons for RT

Session #	Model	Fixed Effects	Random Effects	AIC	Δ AIC	-2LL Statistic	*p*
1	Model 1	None	By-Subject random-intercepts	84395.91	-42194.96		
	Model 2	Model 1 + Group	Same as Model 1	84397.90	-42194.95	$X^2(1) = .01$	.91
	Model 3	Model 2 + Item	Same as Model 1	84352.44	-42171.22	$X^2(1) = 47.46$	<.001
	Model 4	Same as Model 3	Model 1 + by-Item random slope	84266.31	-42126.15	$X^2(2) = 90.14$	<.001
	Model 5	Same as Model 4 + Verb	Same as Model 4	84267.99	-42125.99	$X^2(1) = .32$	.57
	Model 6	Model 5 + Group x Item interaction	Same as Model 4	84249.91	-42115.96	$X^2(1) = 20.08$	< .001
	Model 7	Model 4 + Group x Item interaction	Same as Model 4	84248.22	-42116.11	$X^2(1) = 20.08$	<.001
2	Model 1	None	By-Subject random-intercepts	84161.99	-42077.99		
	Model 2	Model 1 + Group	Same as Model 1	84163.08	-42077.54	$X^2(1) = .91$	.34
	Model 3	Model 2 + Item	Same as Model 1	84140.87	-42065.44	$X^2(1) = 24.20$	<.001
	Model 4	Same as Model 3	Model 1 + by-Item random slope	84116.50	-42051.25	$X^2(2) = 28.38$	<.001
	Model 5	Same as Model 4 + Verb	Same as Model 4	84116.87	-42050.44	$X^2(1) = 1.62$	.20
	Model 6	Model 5 + Group x Item interaction	Same as Model 4	84103.17	-42042.59	$X^2(1) = 15.69$	<.001
	Model 7	Model 4 + Group x Item interaction	Same as Model 4	84102.8	-42043.40	$X^2(1) = .15.69$	<.001
3	Model 1	None	By-Subject random-intercepts	125492.0	-62743		
	Model 2	Model 1 + Group	Same as Model 1	125490.6	-62741.30	$X^2(1) = 3.39$	.07
	Model 3	Model 2 + Item	Same as Model 1	125331.9	-62660.94	$X^2(1) = 160.72$	< .001
	Model 4	Same as Model 3	Model 1 + by-Item random slope	125190.9	-62588.47	$X^2(2) = 144.93$	< .001
	Model 5	Same as Model 4 + Verb	Same as Model 4	125192	-62588.23	$X^2(1) = .49$	.48
	Model 6	Model 5 + Group x Item interaction	Same as Model 4	125176.1	-62579.05	$X^2(1) = 18.35$	< .001
	Model 7	Model 4 + Group x Item interaction	Same as Model 4	125174	-62579.29	$X^2(1) = 18.36$	< .001
4	Model 1	None	By-Subject random-intercepts	193608.5	-96801.23		
	Model 2	Model 1 + Group	Same as Model 1	193609.1	-96800.55	$X^2(1) = 1.37$	.24
	Model 3	Model 2 + Item	Same as Model 1	193357.0	-96673.51	$X^2(1) = 254.07$	< .001
	Model 4	Same as Model 3	Model 1 + by-Item random slope	192961.5	-96473.75	$X^2(1) = 399.53$	< .001
	Model 5	Same as Model 4 + Verb	Same as Model 4	192951.5	-96467.74	$X^2(1) = 12.02$	.001
	Model 6	Model 5 + Group x Item interaction	Same as Model 4	192927.6	-96454.78	$X^2(1) = 25.91$	< .001
	Model 7	Model 4 + Group x Item interaction	Same as Model 4	192937.5	-96460.76	$X^2(1) = 11.95$	.001

Note. Grey shading indicates optimal model

Table 4.
Summary of fixed effects for accuracy

Session #	Parameter	Estimate	95% CIs for Estimate	SE	z-value	*p*	R^2
1	(intercept)	2.13	1.15, 3.10	.49	4.28	< .001	.01
	Group	.49	.04, .94	.23	2.13	.03	.01
	Item	.07	.04, .09	.02	4.23	< .001	.01
	Verb	.19	-.26, .65	.23	.83	.40	.01
	Group X Item	-.02	-.03, -.01	.01	-3.12	.002	.01
2	(intercept)	2.34	1.05, 3.63	.66	3.56	<. 001	.02
	Group	.74	.15, 1.32	.29	2.47	.01	.02
	Item	.05	.02, .08	.01	3.68	<.001	.02
	Verb	.13	-.28, .54	.21	.63	.53	.02
	Group X Item	.02	-.03, -.01	.01	-3.77	< .001	.02
3	(intercept)	2.15	1.32, 2.98	.42	5.07	<. 001	.02
	Group	.46	.09, .84	.19	2.42	.02	.02
	Item	.04	.03, .06	.01	5.33	<. 001	.02
	Verb	-15	-.45, .15	.15	-.96	.34	.02
	Group X Item	-.01	-.02, -.01	.00	-4.13	<. 001	.02
4	(intercept)	.45	-.06, .96	.26	1.73	.08	.04
	Group	.91	.67, 1.15	.12	7.42	<. 001	.04
	Item	.02	.01, .20	.00	8.85	<. 001	.04
	Verb	.06	-.02, .03	.07	.82	.41	.04
	Group X Item	-.01	-.01, -.01	.00	-7.9	< .001	.04

Table 5.
Summary of fixed effects for RT

Session #	Parameter	Estimate	95% CIs for Estimate	SE	DF	t-value	*p*	R^2
1	(intercept)	1912.78	1660.43, 2165.14	128.77	5033	14.85	<.001	.09
	Group	-237.90	-355.49, -120.32	58.59	51	-4.06	<.001	.09
	Item	-13.13	-17.33, -8.94	2.14	5033	-6.14	<.001	.09
	Group X Item	4.81	2.90, 6.72	.97	5033	4.94	<.001	.09
2	(intercept)	1967.85	1700.78, 2234.91	136.28	4843	14.44	<.001	.03
	Group	-260.55	-387.28, -122.83	63.09	49	-4.13	<.001	.03
	Item	-12.70	-17.35, -8.06	2.37	4843	-5.36	<.001	.03
	Group X Item	4.70	2.55, 6.85	1.09	4843	4.28	<.001	.03
3	(intercept)	2353.70	2121.63, 2585.77	118.42	7434	19.88	<.001	.09
	Group	-298.57	-406.59, -190.56	53.79	50	-5.55	<.001	.10
	Item	-10.08	-12.97, -7.19	1.48	7434	-6.83	<.001	.09
	Group X Item	3.15	1.83, 4.46	.67	4.69	4.69	<.001	.09
4	(intercept)	2576.59	2307.74, 2845.43	137.18	11392	18.78	<.001	.11
	Group	-376.77	-501.73, -251.82	62.25	51	-6.05	<.001	.12
	Item	-8.82	-11.06, -6.58	1.14	11392	-7.72	<.001	.11
	Verb	-70.69	-110.76, -30.63	20.44	11392	-3.46	.001	.11
	Group X Item	3.00	1.99, 4.02	0.52	11392	5.78	<.001	.11

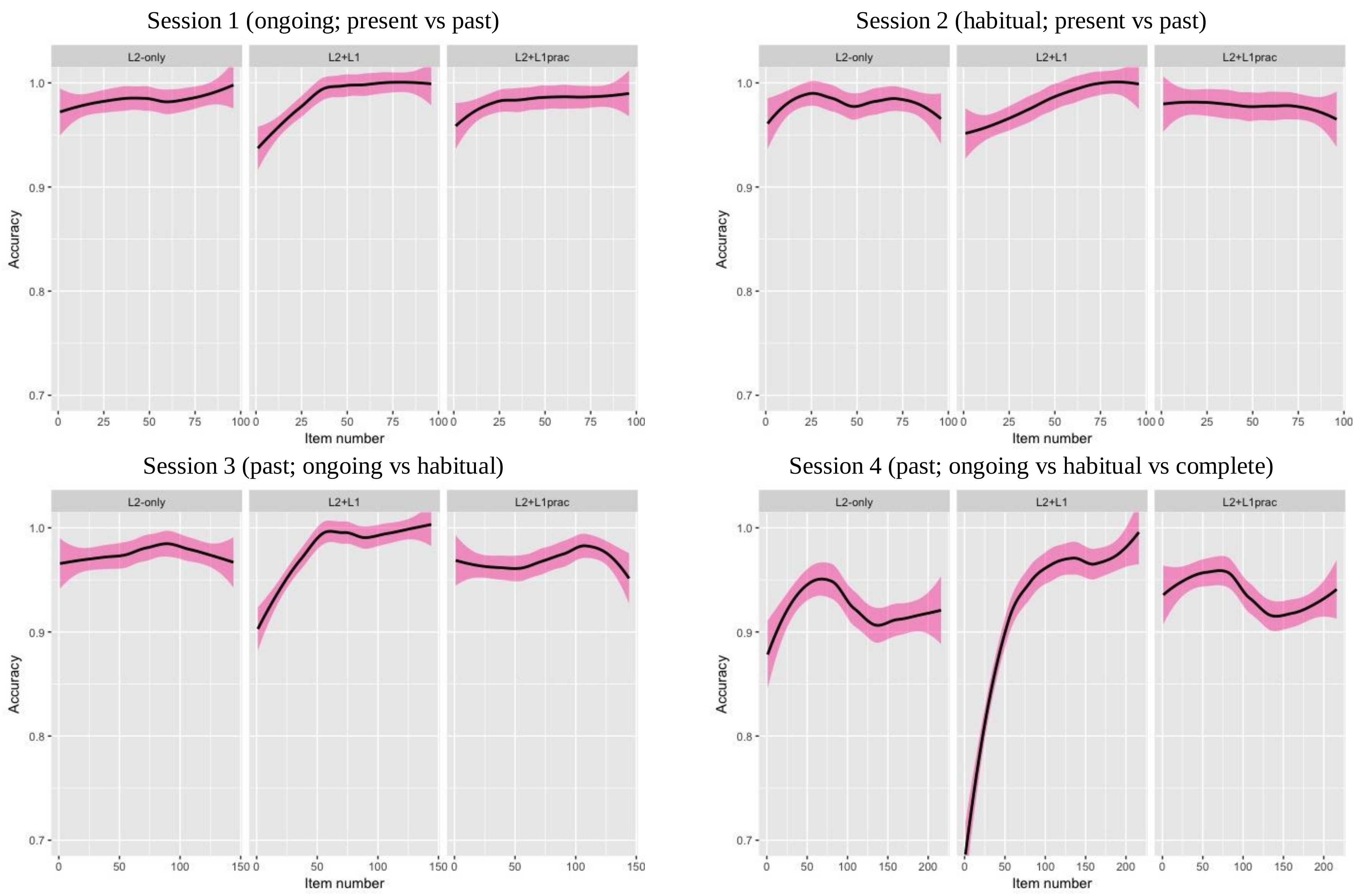
Table 6.
Linear regression results for CV scores in each practice session for the L2+L1 group

	Session 1	Session 2	Session 3	Session 4
Intercept	.64***	.61**	.71***	.74***
[95% CIs]	[.54, .75]	[.54, .67]	[.67, .75]	[.71, .77]
(SE)	(.05)	(.03)	(.02)	(.01)
Item	.000	.002**	-.001***	-.001***
[95% CIs]	[-.00, .00]	[.00, .00]	[-.00, .00]	[-.00, -.00]
(SE)	(.00)	(.00)	(.000)	(.00)
F	(1, 94) = .01	(1, 94) = 9.6	(1, 142) = 47.67	(1, 214) = 102
p	.91	.003	<. 001	<. 001
R^2	.00	.09	.25	.32

Note. *** = *p* <. 001, ** = *p* <.01

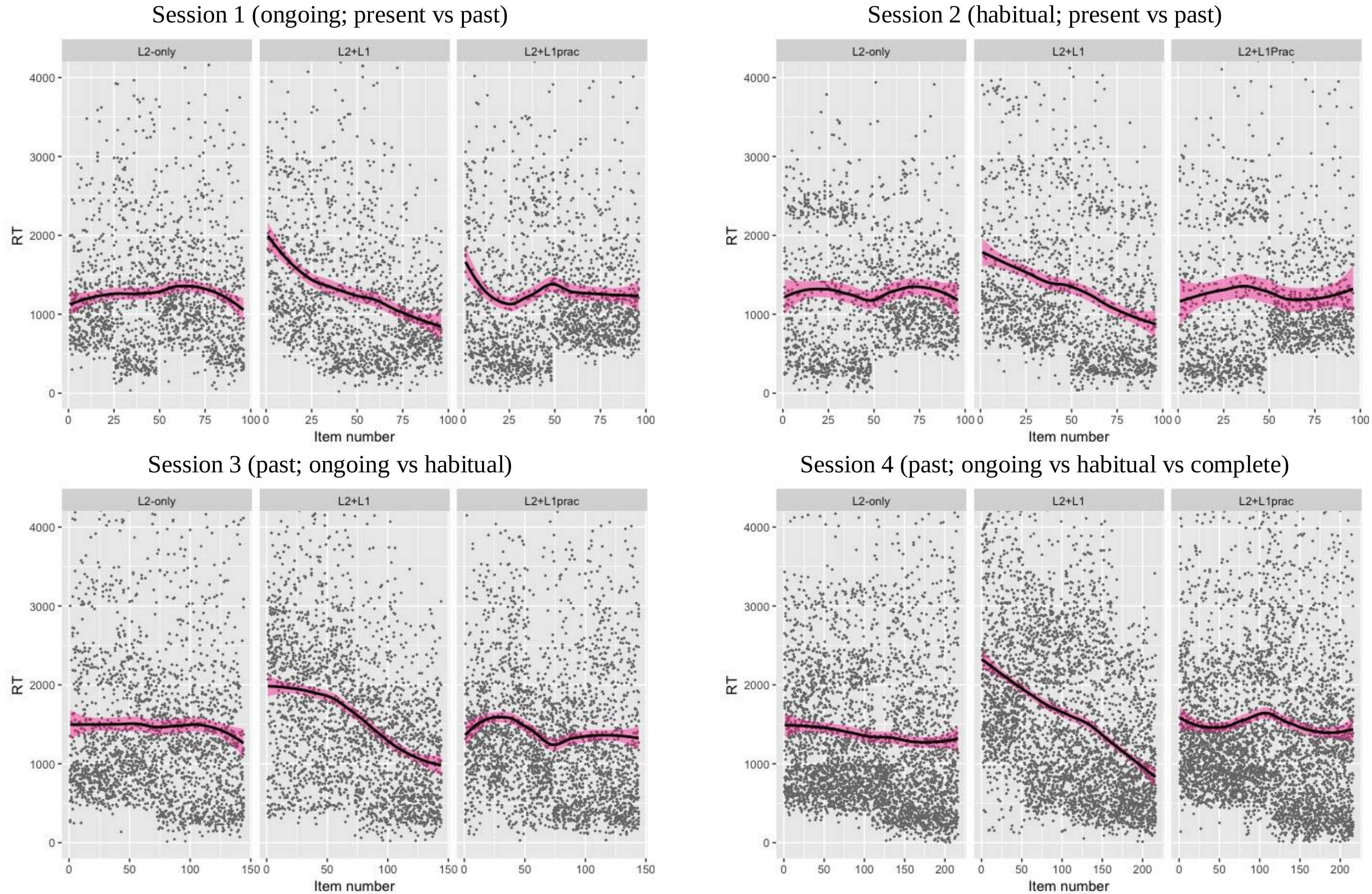
FIGURES

Figure 1. Accuracy scores over time in each training session (black line = regression line, pink shading = 95% CIs)[6]



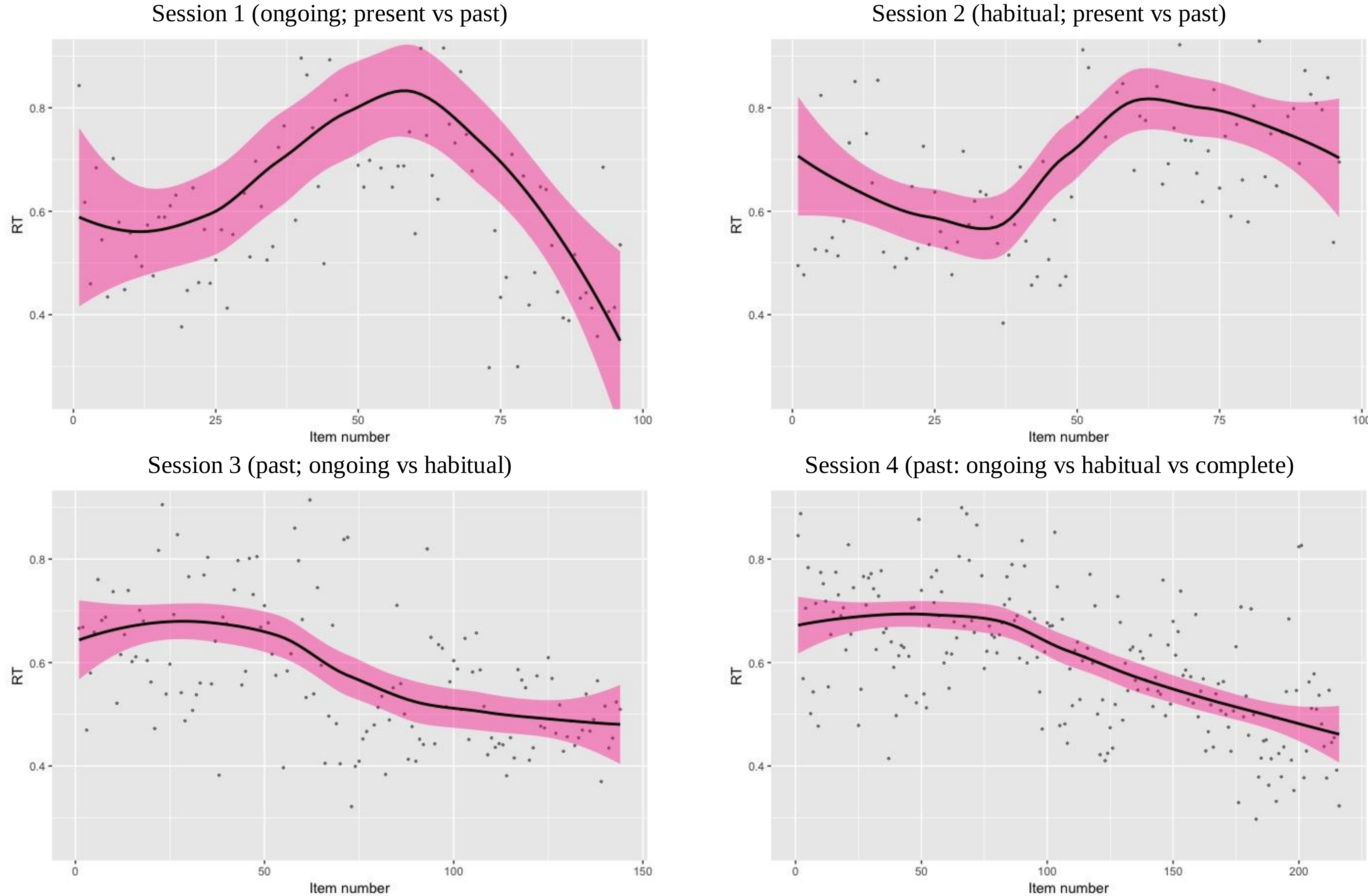
FIGURES

Figure 2. Reaction times over time in each training session (black line = regression line, pink shading = 95% CIs, grey dots = individual data points)



FIGURES

Figure 3. CVs over time for L2+L1 group in all sessions (black line = regression line, pink shading = 95% CIs, grey dots = individual data points)



FIGURES

APPENDIX A

Table A1.
Frequencies of French stimuli used all treatments

		Listening	Reading	Total
IMP Ongoing/Interrupted	Session 1	24	24	48
	Session 3	36	36	72
	Session 4	36	36	72
IMP Habitual	Session 2	24	24	48
	Session 3	36	36	72
	Session 4	36	36	72
	TOTAL IMP	192	192	384
IMP juxtaposed with…				
Présent	Session 1	24	24	48
Présent	Session 2	24	24	48
Passé Composé	Session 4	36	36	72
	GRAND TOTALS	276	276	552

Table A2.
Frequencies of EI received during training when selecting incorrect answers, by group

	L2+L1 (n=17)	L2-only (n=17)	L2+L1prac (n=19)
Session 1: Ongoingness, past vs. present			
After incorrectly responding ‘MAINTENANT’	9	5	13
After incorrectly responding ‘DANS LE PASSÉ’	8	10	10
Session 2: Habituality, past vs. present			
After incorrectly responding ‘MAINTENANT’	23	21	29
After incorrectly responding ‘DANS LE PASSÉ’	7	9	7
Session 3: Ongoing vs. Habitual, past only			
After incorrectly responding ‘ONGOING / INTERRUPTED’	60	63	57
After incorrectly responding ‘REGULARLY REPEATED'	25	28	26
Session 4: Ongoing vs. Habitual vs. Complete, past only			
After incorrectly responding ‘ONGOING / INTERRUPTED’	111	111	108
After incorrectly responding ‘REGULARLY REPEATED'	71	77	68
After incorrectly responding ‘COMPLETE'	87	93	83

Table A3.
Frequencies of English stimuli used in L2+L1 and L2+L1prac treatments

		Listening	Reading	Total
Past Progressive (Ongoing)	Session 1	8	8	16
	Session 3	12	12	24
	Session 4	8	8	16
Past Simple (Habitual)	Session 2	8	8	16
	Session 3	12	12	24
	Session 4	8	8	16
TOTAL ONGOING & HABITUAL		56	56	112
Juxtaposed with…				
Present Progressive	Session 1	8	8	16
Present Simple	Session 2	8	8	16
Past Simple	Session 4	8	8	16
GRAND TOTALS		80	80	160

APPENDIX B

Examples of French stimuli

Session 1. Ongoing: Participants choose between past vs. present

Condition	*Statives*	*Activities*	*Accomplishments*	*Achievements*
	Elle… 'she…'	Elle…	Elle…	Elle…
Past (IMP)	habitait tout seul 'was living alone'	fumait des cigares 'was smoking cigars'	écrivait une lettre 'was writing a letter'	finissait ses devoirs 'was finishing her homework'
Present (PRES)	adore la musique française 'is loving the French music'	nage dans la piscine 'is swimming in the pool'	chante une chanson 'is singing a song'	frappe son ami 'is hitting her friend'

Note. IMP=Imparfait, PRES= Présent

Session 2. Habitual: Participants had to choose between past vs. present

Condition	*Statives*	*Activities*	*Accomplishments*	*Achievements*
	Il… 'he…'	Il…	Il…	Il…
Past (IMP)	portait une cravate 'wore/used to wear a tie'	conduisait avec la famille 'drove/used to drive with his family'	regardait un film 'watched/used to watch a film'	perdait sa montre 'lost/used to lose his watch'
Present (PRES)	adore la musique française 'loves French music'	boit du café 'drink coffee'	joue un match de foot 'plays a game of football'	trouve sa voiture 'finds his car'

Note. IMP=Imparfait, PRES= Présent

Session 3. Past: Participants had to choose between ongoing vs. habitual (‘regularly repeated’).

N.B. Words are underlined for illustrative purposes only, to indicate which verb the participants had to respond to. Only main->subordinate clause ordering is illustrated here.

Condition	*Statives*	*Activities*	*Accomplishments*	*Achievements*
	Elle… ‘she…’	Elle…	Elle…	Elle…
Ongoing (IMP +PC)	aimait les fleurs quand son enfant a commencé à pleurer ‘was liking the flowers when her child began to cry’	mangeait un sandwich quand la cloche a sonné ‘was eating a sandwich when the bell rang’	lisait le journal quand son chef a sonné à la porte ‘was reading the paper when her boss rang the doorbell’	quittait la maison quand son ami l’a appelé ‘was leaving the house when her friend called her’
Habitual (IMP +IMP)	savait manger sainement quand elle allait à la gym ‘knew/used to know how to eat healthily when she went to the gym’	conduisait avec la famille quand elle habitait avec son mari ‘drove/used to drive with the family when she lived with her husband’	jouait un match de tennis quand il allait à la gym ‘played/used to play a game of tennis when he went to the gym’	remarquait les touristes quand elle habitait à Paris ‘noticed/used to notice the tourists when she lived in Paris’

Note. IMP=Imparfait, *PC= Passé Composé*

Session 4. Past: Participants had to choose between ongoing vs. habitual (regularly repeated) vs. complete.

N.B. Words are underlined for illustrative purposes only, to indicate which verb the participants had to respond to. Only main->subordinate clause ordering is illustrated here.

Condition	*Statives*	*Activities*	*Accomplishments*	*Achievements*
	Il… 'he…'	Il…	Il…	Il…
Ongoing (IMP + PC)	adorait son nouvel album quand il est allé au travail 'was loving his new album when he went to work'	fumait des cigares quand sa femme est arrivée 'was smoking cigars when his wife arrived'	mangeait deux pommes quand le professeur est arrivé 'was eating two apples when the teacher arrived'	finissait son petit-déjeuner quand la cloche a sonné 'was finishing his breakfast when the bell rang'
Habitual (IMP + IMP)	portait une cravate quand il allait à l'école 'wore/used to wear a tie when he went to school'	rigolait dans le bar quand il buvait avec ses amis 'laughed/used to laugh in the bar when he drank with his friends'	nageait deux mètres quand il habitait près de la mer 'swam/used to swim two metres when he lived by the sea'	trouvait ses clés quand il passait le weekend chez lui 'found/used to find his keys when he spent the weekend at his place'
Complete (PC + PC)	a aimé la peinture quand sa femme est arrivée 'loved the painting when his wife arrived	a conduit sa voiture quand il a reçu un SMS 'drove his car when he received a text message'	a lu un livre quand sa femme a commencé à jouer au piano 'read a book when his wife started to play the piano'	est sorti de la maison quand il a commencé à pleuvoir 'left the house when it started to rain'

Note. IMP=Imparfait, *PC= Passé Composé*

Examples of English stimuli

Session 1. Ongoing: Participants had to choose between past vs. present.

Condition	*Statives*	*Activities*	*Accomplishments*	*Achievements*
	She…	She…	She…	She…
Past (Past Progressive	was hating all the noise	was playing cards	was eating a sandwich	was ringing his friend
Present (Present Progressive)	is enjoying the weather	is listening to the music	is walking to the stage	is finishing his drink

Session 2. Habitual: Participants had to choose between past vs. present.

Condition	*Statives*	*Activities*	*Accomplishments*	*Achievements*
	He…	He…	He…	He…
Past (Past Simple)	liked the weather	did the washing up	ran to the park	found his watch
Present (Present Simple)	adores his boat	reads in the park	drinks a glass of wine	hits the wall

Session 3. Past: Participants had to choose between ongoing vs. habitual.

N.B. Words are underlined for illustrative purposes only, to indicate which verb the participants had to respond to. Only main->subordinate clause ordering is illustrated here.

Condition	*Statives*	*Activities*	*Accomplishments*	*Achievements*
	She…	She…	She…	She…
Ongoing (Past Progressive)	<u>was listening</u> to music when her phone rang	<u>was smoking</u> when the bus arrived	<u>was drinking</u> a cup of tea when the cup broke	<u>was knocking</u> at the door when her phone rang
Habitual (Past Simple)	<u>enjoyed</u> the weekends when she didn't work	<u>spoke</u> French when she had a French boyfriend	<u>read</u> the newspaper when she had the time	<u>arrived</u> on time when she lived closer to work

Session 4. Past: Participants had to choose between ongoing vs. habitual vs. complete.

N.B. Words are underlined for illustrative purposes only, to indicate which verb the participants had to respond to.

Condition	*Statives*	*Activities*	*Accomplishments*	*Achievements*
	He…	He…	He…	He…
Ongoing (Past Progressive)	was hating the soup when the waiter arrived	was reading when the baby started to cry	was playing a game of football when he fell	was arriving home when his phone rang
Habitual (Past Simple)	liked museums when he used to go on holiday with his dad	listened to music when he used to live alone	ate an apple when he used to make his own lunch	left the house when the postman used to knock
Complete (Past Simple)	knew the answer when the teacher started to ask questions	cycled to work when it started to rain	wrote a few sentences when the pen broke	finished the race when his wife called

APPENDIX C

Ongoingness (Present vs. Past), Session 1

<table>
<tr><th></th><th>L2-only</th><th>L2+L1</th></tr>
<tr><td>Pre-practice EI</td><td>[A six-second video clip of man eating an apple. The apple was never fully eaten.]

To describe this you could say:

Il mange une pomme
Or

Il mangeait une pomme

The difference between these two is:
Il mange = ongoing action RIGHT NOW
Il mangeait = ongoing action IN THE PAST

The ends of the verbs distinguish between an ongoing action in the present versus past e.g. [Four verbs presented in pairs, aurally and in writing]:
<table><tr><th>Présent
RIGHT NOW</th><th>Imparfait
IN PAST</th></tr><tr><td>regarde
[ʀəgaʀd]</td><td>regardait
[ʀəgaʀdɛ]</td></tr></table></td><td>[Same video again as L2-only treatment]

To describe this you could say:

He is eating an apple
Or
He was eating an apple

The difference between these two is:
‘he is eating’ = ongoing action RIGHT NOW
‘he was eating’ = ongoing action IN PAST”

To identify ongoing meaning in the present versus the past, you need to focus on the auxiliary.

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).”</td></tr>
<tr><td>Practice</td><td>96 (48 listening; 48 reading) items. Learners must identify whether an ongoing event is taking place “MAINTENANT” (right now) or “DANS LE PASSÉ” (in the past), e.g.:
Il…
(1) fait du shopping (‘is shopping’)
(2) parlait français (‘was speaking French’)</td><td>32 (16 listening; 16 reading) items. Learners must identify whether an ongoing event is taking place “RIGHT NOW” or “IN THE PAST”, e.g.
He…
(1) is eating a sandwich
(2) was running to the shop</td></tr>
</table>

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 thing: *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*”