

Bilingual and monolingual adults learning an additional language: ERPs reveal differences in syntactic processing*

SARAH GREY

Department of Modern Languages and Literatures, Fordham University

CRISTINA SANZ

Department of Spanish and Portuguese, Georgetown University

KARA MORGAN-SHORT

Department of Hispanic and Italian Studies, Department of Psychology, University of Illinois-Chicago

MICHAEL T. ULLMAN

Department of Neuroscience, Georgetown University

(Received: September 03, 2016; final revision received: June 29, 2017; accepted: July 26, 2017)

It has been suggested that bilinguals learn additional languages ‘better’ than monolinguals. However, evidence is sparse, particularly for grammar. We examined behavioral and neural correlates of learning an additional (artificial) language in early Mandarin–English bilinguals, compared to English monolinguals. Following grammar instruction, participants practiced comprehension and production, and judged grammaticality at low and high proficiency while event-related potentials (ERPs) were acquired. Bilinguals and monolinguals did not differ on behavioral measures, but showed distinct ERP patterns. At low proficiency only bilinguals showed a P600, a common ERP correlate of syntactic processing in native speakers of languages. At high proficiency both groups showed P600s, though the monolinguals also evidenced an anterior positivity not typically found in native speakers of languages during syntactic processing. These findings suggest that, even without bilingual/monolingual behavioral differences, bilinguals show ERP patterns for an additional language that are more similar to those of native speakers of languages.

Keywords: grammatical processing, ERPs, syntax, language learning

Introduction

It is often claimed that bilinguals can learn a third language (L3) ‘better’ than monolinguals learn a second language (L2) (Abu-Rabia & Sanitsky, 2010; Cenoz, 2003). Indeed, a number of studies support this claim. For instance, adult bilinguals have shown superior vocabulary learning and processing in an additional language as compared to monolinguals (Antoniou, Liang, Ettliger & Wong, 2015; Bartolotti & Marian, 2012; Bartolotti, Marian, Schroeder & Shook, 2011; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a, 2009b; Kaushanskaya & Rehtzigel, 2012; Keshavarz & Astaneh, 2004; Wang & Saffran, 2014; Zare & Mobarakeh, 2013), as well as superior pragmatic skill development (Safont Jorda, 2003) and better general additional language proficiency (Cenoz & Valencia, 1994; Sanz, 2000). The benefits of bilingualism with respect

to language learning have been attributed to bilinguals’ greater metalinguistic awareness (Dillon, 2009; Jessner, 2008), better use of learning strategies (Kemp, 2007), the availability of a broader linguistic repertoire (Cenoz, 2013; De Angelis, 2007), and better management of cross-language interference (Bartolotti & Marian, 2012), as compared to monolinguals.

Although quite a few studies suggest bilingual advantages for the functions listed above, others show no group differences (or show an advantage for monolinguals, Okita & Jun Hai, 2001), both for lexical learning (Antoniou et al., 2015; Gonzalez-Ardeo, 2000) and for preposition learning (Gibson, Hufeisen & Libben, 2001). Thus, clear conclusions are difficult to draw and more research is needed, especially with regard to the linguistic domain of grammar, which has been addressed by only three studies (Cox, 2017; Nation & McLaughlin, 1986; Nayak, Hansen, Krueger & McLaughlin, 1990). Moreover, weaknesses and gaps within these studies¹, as well as heterogeneity across them, make the existing literature on grammar even more inconclusive. The available research, then, provides a remarkably narrow

* This research was supported by NSF-BCS 1124144 to Sarah Grey, Cristina Sanz, and Michael T. Ullman and a *Language Learning* journal dissertation grant to Sarah Grey. Portions of this research were presented at the annual meetings of the American Association for Applied Linguistics and the Cognitive Neuroscience Society in 2014. We thank three anonymous reviewers for their valuable feedback. We especially thank Janire Zalbidea for assistance with data collection for the bilingual group.

Address for correspondence:

Sarah Grey, Department of Modern Languages and Literatures, Faber Hall 556, Fordham University, Bronx, NY 10458
sgrey4@fordham.edu

¹ There is L3 research on the influence of linguistic transfer (e.g., Falk & Bardel, 2010; Rah, 2010; Rothman, Iverson, & Judy, 2011; Sanz, Park, & Lado, 2015), but such work focuses on the cross-linguistic influences of first, second, and third language transfer, and does not address bilingual/monolingual differences in additional language learning/processing. Therefore it is not included here.

empirical basis on which to make claims about bilingual advantages in language learning, particularly grammar learning – that is, rule-governed aspects of syntax (e.g., word order) as well as morphologically-marked features such as grammatical gender agreement or case-marking. Examining grammar is important not only because it underlies the flexibility and creativity of language, but also because neurocognitive theories of additional language learning posit differential levels of attainment for aspects of grammatical as compared to lexical learning and processing (e.g., Clahsen & Felser, 2006a, 2006b; Paradis, 2009; Ullman, 2005, 2015). Additionally, the existing research on grammar has examined only behavioral measures of learning, so it is unknown how bilinguals and monolinguals compare on brain measures of grammar processing. In this study, we investigate these issues by examining both behavioral and brain (electrophysiological) measures of potential bilingual/monolingual differences in adult additional language learning of grammar.

Bilingual L3 versus monolingual L2 grammar learning

As mentioned above, to our knowledge there are only three published studies on adult language learning that have investigated potential differences in grammar learning between monolinguals and either bilinguals or multilinguals (Cox, 2017; Nation & McLaughlin, 1986; Nayak et al., 1990).

An early study compared monolinguals, bilinguals, and multilinguals (with mixed language backgrounds) on their learning of an artificial grammar (Nation & McLaughlin, 1986); that is, a system of elements, such as consonant letter strings, whose structure is determined by a set of rules. All learners were exposed to the same visual input (20 exemplar strings of letters) under either an implicit or explicit condition. Learners in the ‘implicit’ condition were given no specific instructions to learn the grammar rule, whereas those in the ‘explicit’ condition were informed that the system was rule-based and they should try to discover the rule. The results from a grammaticality judgment test following training showed that under the implicit condition the multilinguals outperformed both the bilinguals and monolinguals (who did not differ from each other), whereas the three groups did not differ in the explicit rule-search condition. The authors suggested that multilinguals are better able to allocate processing resources to generate rules under implicit conditions. Although artificial grammar paradigms are designed to elucidate aspects of the human capacity to learn language-like rules (Gomez, 1997; Reber, 1967, 1969), they lack the characteristic form-meaning mappings of natural languages. It is therefore uncertain whether and how results from artificial

grammar paradigms can be generalized to natural languages.

Subsequently, Nayak et al. (1990) examined monolinguals and multilinguals (mixed language backgrounds; no bilinguals were tested) in their learning of an artificial language: that is, a system of elements whose structure is determined by complex grammar rules, and whose words and grammar are characterized by form-meaning mappings. In this study, the language consisted of 40 visually-presented sentences, and the words in its lexicon were mapped to specific geometric figures (e.g., a rectangle). Both multilinguals and monolinguals were trained under one of two conditions, which the authors referred to as the ‘memory’ and ‘rule-discovery’ conditions. Learners in the memory condition were asked to memorize the phrases, whereas those in the rule-search condition were instructed to determine the rules that governed word order in the language. After the training phase, learners completed a vocabulary test as well as a grammaticality judgment test that measured word-order learning. The results showed no monolingual/multilingual group differences for vocabulary learning under either condition. For word order, the multilinguals outperformed the monolinguals in the rule-search condition but not in the memory condition. These results were similar to those reported by Nation and McLaughlin (1986) in that both studies reported advantages of multilinguals over monolinguals for grammar learning. However, the two studies differed in that these advantages were found by Nation and McLaughlin (1986) in the ‘implicit’ condition, but by Nayak et al. (1990) in the rule-search (explicit) condition. Nayak et al. attributed this difference between the studies to the fact that their artificial language was more complex than the artificial grammar used by Nation and McLaughlin, and therefore required more linguistic processing than the pattern recognition processes examined by Nation and McLaughlin. Additionally, the studies differed as to what characterized their implicit conditions (no instructions in Nation et al., 1986 vs. instructions to memorize in Nayak et al., 1990), which further complicates direct comparisons of the studies’ outcomes for multilingual/monolingual performance.

The third study was conducted in the context of the Latin Project, which was designed by Cristina Sanz and colleagues to examine bilingualism and the effects of both explicit instruction and feedback in the L2/L3 learning of morphosyntactic case, specifically agent/patient functions (e.g., Cox & Sanz, 2015; Cox, 2017; Lado, 2016; Stafford, Bowden & Sanz, 2012). The project employs a miniature version of Latin (a dead language). Relevant to the present study, Cox (2017) tested English L1–Spanish L2 bilinguals and English monolinguals (all > 60 years old) on the learning of Latin morphosyntactic case-function assignment – under

either a visually-presented ‘explicit instruction’ condition, that provided metalinguistic grammar explanations, or a condition that was less explicit, in that it did not provide metalinguistic explanations (audio-visual presentation of meaningful sentences and pictures only). Cox (2017) administered four assessments of learning: written sentence interpretation, aural sentence interpretation, grammaticality judgment, and written sentence production. The results revealed that the bilinguals outperformed the monolinguals on the sentence interpretation tasks (which depended on case-assignment), regardless of instructional condition. Additionally, for grammaticality judgment and written sentence production, the results showed a statistical trend for the bilinguals in the explicit instruction condition to outperform monolinguals who did not receive metalinguistic explanations; whereas there were no significant differences between bilinguals or monolinguals who received explicit instruction. Although this study provides evidence for superior learning of additional language grammar in bilinguals compared to monolinguals under some conditions, it focused on older adults (> 60 years old), which limits the study’s comparability to research on younger adult bilingual and monolingual learners.

Although all three studies are informative, an important advantage of the Nayak et al. (1990) study is its use of a well-controlled artificial language paradigm. Artificial languages, in contrast to artificial grammars, provide learners with form-meaning mappings, and are thus more similar to natural languages. Additionally, their rules are typically more complex than those of artificial grammars. Artificial languages also have certain advantages over full natural languages, as well as over miniature languages based on real languages such as Latin. In particular, artificial languages limit some of the major confounds and weaknesses inherent in natural (and miniature) language research of adult language learning. For example, in natural language research, learners often have uncontrolled or unknown types and amounts of language exposure. Moreover, in both full natural language and miniature language research, language transfer issues are often not fully controlled for (but see Tagarelli, 2014). Overall, these problems make it difficult to clearly elucidate issues of interest in language learning. Finally, learning a full natural language to high levels of proficiency takes several (often many) years, which makes it extremely difficult to examine the trajectory of learning.

Artificial languages are less susceptible to these confounds and weaknesses. First, both the amount and type of language exposure can be fully controlled for (Friederici, Steinhauer & Pfeifer, 2002b; Morgan-Short, Sanz, Steinhauer & Ullman, 2010; Morgan-Short, Steinhauer, Sanz & Ullman, 2012b). Additionally, grammatical and lexical similarity to already-known

languages can be carefully manipulated, allowing one to control for issues of language transfer (Sanz & Lado, 2008; Lado & Sanz, 2016; note that transfer cannot be well controlled with artificial grammars, since they lack form-meaning connections). Another advantage of artificial languages is that they can generally be learned to high proficiency in hours to days, which provides the means to closely examine the trajectory of language learning (Morgan-Short et al., 2010; Morgan-Short et al., 2012b). Importantly, the neural correlates of artificial language learning and processing are very similar to those of natural languages (Friederici et al., 2002b; Morgan-Short et al., 2010; Morgan-Short, Finger, Grey & Ullman, 2012a; Morgan-Short et al., 2012b; Morgan-Short, Deng, Brill-Schuetz, Faretta-Stutenberg, Wong & Wong, 2015). This suggests that artificial language paradigms are reasonable models of natural language learning. In fact, artificial language learning performance has been shown to correlate positively with natural L2 learning performance (Ettliger, Morgan-Short, Faretta-Stutenberg & Wong, 2015). Indeed, there is a robust body of research that has used artificial language systems to investigate compelling questions in adult language learning and processing (DeKeyser, 1996, 1997; Friederici et al., 2002b; Grey, Williams & Rebuschat, 2014, 2015; Hama & Leow, 2010; Morgan-Short et al., 2010; Morgan-Short et al., 2012a; Morgan-Short et al., 2012b; Morgan-Short et al., 2015; Rebuschat & Williams, 2012; Williams, 2005). Thus, in the present study, we employ an artificial language design.

Gaps and limitations in existing L3/L2 research

Overall, the available evidence offers quite limited insight into the potential effects of bilingualism on L3 learning of grammar, compared to monolingual L2 learning. As we have seen, early work using an artificial grammar paradigm found superior performance in multilinguals compared to bilinguals and monolinguals, with no differences between the latter two (Nation & McLaughlin, 1986). However, caution is warranted in generalizing findings from artificial grammar learning to natural language learning. A subsequent study, which used an artificial language, showed evidence of superior performance in multilinguals as compared to monolinguals on syntactic word order (Nayak et al., 1990). However, this study did not test bilinguals. More recent work with a miniature language found that bilinguals outperformed monolinguals at sentence interpretation tasks that depended on morphosyntactic knowledge (Cox, 2017), but the study included only older adults.

Additionally, none of these studies examined bilinguals and monolinguals longitudinally over the course of learning, from lower to higher levels of proficiency – even though such an approach is critical for revealing any L3/L2 differences in the trajectory of learning, or any differences

that may be limited to particular points along the learning trajectory. Finally, previous research has examined only behavioral measures of learning, so possible neural correlates of any hypothesized bilingual advantages in language learning remain unclear, ultimately limiting the depth of our understanding of how bilingual L3 and monolingual L2 learning might differ.

ERPs and adult language learning

In the present study, both behavioral and event-related potential (ERP) measures were acquired to examine the neurocognition of language processing at low and high additional language proficiency, in bilingual and monolingual language learners. ERPs offer several benefits for studying language learning and processing. First, the ERP components that are typically elicited in both first and second language experiments have been very well studied, and are relatively well understood (Kaan, 2007; Morgan-Short & Tanner, 2014; Steinhauer & Connolly, 2008; Van Hell & Tokowicz, 2010). Second, ERPs can reveal group differences that may not be evident in behavioral measures of performance (e.g., Grey, Tanner & Van Hell, 2017; Tokowicz & MacWhinney, 2005). Third, ERPs can elucidate qualitative differences – as revealed by different components – in how linguistic information is processed, including in the processing of lexical and grammatical features of an L2 (Batterink & Neville, 2013; Foucart & Frenck-Mestre, 2012; Gillon Dowens, Guo, Guo, Barber & Carreiras, 2011; McLaughlin, Osterhout & Kim, 2004; Morgan-Short et al., 2010; Morgan-Short et al., 2012b; Tanner, Inoue & Osterhout, 2014; Tokowicz & MacWhinney, 2005). Furthermore, ERPs can reveal quantitative differences between groups and conditions in factors such as the timing, distribution, and amplitude of components.

ERPs are especially useful for investigating potential bilingual/monolingual differences in adult language learning of grammar, given that there is a well-studied ERP component that has shown reliable associations with grammatical processing: the P600. The P600 is a positive-going waveform with a centro-posterior scalp distribution. Nearly three decades of research have revealed that the P600 is remarkably sensitive to and reliably elicited by (morpho)syntactic violations, for example by violations of word order (e.g., Friederici & Mecklinger, 1996), subject-verb agreement (e.g., Osterhout & Mobley, 1995), and nominal (gender/number) agreement (e.g., Barber & Carreiras, 2005). Although there are differences in the specific theoretical descriptions of P600 effects (Chow & Phillips, 2013; Friederici, Hahne & Saddy, 2002a; Osterhout, Kim & Kuperberg, 2012; van de Meerendonk, Kolk, Vissers & Chwilla, 2010), there is a general consensus that P600s reflect the processing of a stimulus in conflict with an expected linguistic representation,

and a late attempt at resolving or reanalyzing this conflict. Note that in many studies, P600s are preceded by frontally-distributed and occasionally left-lateralized anterior negativities, termed LANs, and that this biphasic LAN-P600 pattern is considered by some researchers to be a key neural signature of (morpho)syntactic processing (Friederici et al., 2002a; Molinaro, Barber & Carreiras, 2011). Specifically, the biphasic response is considered to reflect the early, automatic detection (LAN) followed by the later, controlled reanalysis (P600) of (morpho)syntactic violations. However, whereas the P600 is robust and replicable across ERP studies of (morpho)syntactic processing, the LAN is variable and often absent, and its functional significance is still debated (for recent discussions of the LAN, including LAN versus N400 patterns, see Molinaro, Barber, Caffarra & Carreiras, 2015; Steinhauer & Drury, 2012; Tanner, 2015).

P600s are found both in L1 and L2 processing. In L2 processing, P600s are commonly elicited (as in L1) in response to (morpho)syntactic violations, in both L2 classroom and L2 immersion exposure contexts (Batterink & Neville, 2013; Gillon Dowens et al., 2011; Gillon Dowens, Vergara, Barber & Carreiras, 2009; Tanner, McLaughlin, Herschensohn & Osterhout, 2013; White, Genesee & Steinhauer, 2012). P600s in L2 appear to be more reliable at higher than at lower L2 proficiency (for reviews see Bowden, Steinhauer, Sanz & Ullman, 2013; Morgan-Short, 2014; Steinhauer, White & Drury, 2009; Steinhauer, 2014). Additionally, though the biphasic LAN-P600 response has been found in L2 studies, it has ONLY been observed at higher L2 proficiency (e.g., Bowden et al., 2013; Gillon Dowens et al., 2009; Rossi, Hartmüller, Vignotto & Obrig, 2013). Similar to natural second languages, (morpho)syntactic violations in artificial languages also elicit (LAN)-P600 patterns, usually at higher rather than lower proficiency (e.g., Friederici et al., 2002b; Morgan-Short et al., 2010; Morgan-Short et al., 2012a; Morgan-Short et al., 2012b). Thus both P600s and LANs are more commonly observed at higher than lower L2 proficiency.

The observed changes in L2 ERP patterns (for both natural and artificial language) as a function of L2 proficiency have informed current neurocognitive theories of later-learned language processing. The theories, though different in their specific approaches to L1 and L2 processing, generally converge in suggesting that aspects of L2 processing can in principle approximate those of L1 (but see Clahsen & Felser, 2006a, 2006b). In particular, a number of neurocognitive theories propose that L2 grammar can eventually rely on similar mechanisms as L1 when L2 proficiency has reached high enough levels (e.g., Bates & MacWhinney, 1989; Green, 2003; MacWhinney, 2001; Paradis, 2009; Ullman, 2005, 2015).

However, these theoretical approaches have been developed to account primarily for monolingual L2

learning. It remains to be seen whether and how they might be applied to bilingual L3 learning. Indeed, there is very little ERP research that has explicitly studied L3 processing. We are aware of only a few L3 ERP studies, which have examined L1, L2, and L3 word recognition in trilinguals (Aparicio, Midgley, Holcomb, Pu, Lavaur & Grainger, 2012), language inhibition during L1, L2, and L3 digit naming in trilinguals (Guo, Ma & Liu, 2013), or language control in early and late bilinguals' L3 processing (Martin, Strijkers, Santesteban, Escera, Hartsuiker & Costa, 2013). These studies either compared processing within subjects' three languages (Aparicio et al., 2012; Guo et al., 2013), or compared two sets of bilinguals' L3 processing (Martin et al., 2013); that is, none compared trilinguals to bilinguals or monolinguals, or bilinguals to monolinguals in their learning of an additional language. Moreover, these studies examined either lexical processing (Aparicio et al., 2012; Martin et al., 2013) or non-linguistic cognitive control (Guo et al., 2013). To our knowledge no ERP studies on bilingual L3 have examined (morpho)syntactic processing. The present study is well-positioned, then, to be informative not only regarding potential neural correlates of bilingual/monolingual differences in adult language learning, but on bilingual L3 learning more generally.

The current study

In this study we aimed to address gaps and limitations in previous research on potential bilingual advantages in adult additional language learning. Using a well-established artificial language paradigm in which the language is actually spoken and comprehended (Morgan-Short et al., 2015; Morgan-Short, Faretta-Stutenberg, Brill-Schuetz, Carpenter & Wong, 2014; Morgan-Short et al., 2012a; Morgan-Short et al., 2012b; Morgan-Short et al., 2010), we investigated bilingual L3 compared to monolingual L2 grammar processing, specifically syntactic word order. Because previous research suggests that early, relatively balanced bilinguals are most likely to show positive effects of bilingualism on additional language learning (Bartolotti et al., 2011; Cenoz & Valencia, 1994; Kaushanskaya & Marian, 2009a; Sanz, 2000), we recruited only early, relatively balanced bilinguals.

To extend existing behavioral research, we gathered behavioral measures of language comprehension and language production, as well as grammaticality judgments of syntactic word order. To inform the neural bases of bilingual/monolingual adult language learning, we also acquired electrophysiological (ERP) measures of syntactic processing. Moreover, to reveal the longitudinal changes in the behavioral and neural correlates of language learning in bilinguals and monolinguals,

we gathered language comprehension and production measures continuously, and assessed both grammaticality judgments and electrophysiological patterns at two points in the learning trajectory, at low and high proficiency. We investigated the following research questions (RQs):

- RQ 1. Do the bilingual and monolingual groups differ in either their production or comprehension of the additional language?
- RQ 2. Do the bilingual and monolingual groups differ in their grammaticality judgments of syntactic word order of the additional language, at either low or high proficiency?
- RQ 3. Do the two groups differ in their ERP responses to violations of the syntactic word order of the additional language, at either low or high proficiency?

Drawing on prior research that has examined behavioral differences in bilingual/monolingual additional language learning, we predicted that the bilinguals would outperform the monolinguals on comprehension and production (RQ 1) and on grammaticality judgment (RQ 2) – perhaps particularly at low proficiency in the language, since at high proficiency both groups may reach ceiling. Regarding ERP responses (RQ 3), the lack of any prior research on bilingual L3 (morpho)syntactic processing using ERPs precluded strong predictions.

Methods

Participants

Seventeen Mandarin L1–English L2 early bilinguals were tested. These bilinguals were compared to a matched group of native English-speaking monolinguals ($n = 16$) that had been previously tested under the exact same experimental conditions (Morgan-Short et al., 2012b; Morgan-Short et al., 2010). Mandarin–English bilinguals were recruited because Mandarin and English both differ from the artificial language used in the present study, Brocanto2, with respect to their grammatical features, in particular word order (head-first phrase structure in Mandarin and English versus head-final in Brocanto2; see below). This minimizes potential cross-linguistic transfer from the natural language(s) in both groups to Brocanto2, and moreover avoids any transfer differences between the two groups (for transfer considerations in L2/L3 research see Note 1).

Prior to being invited to participate in the study, the bilinguals were screened for inclusionary criteria. Due to the study's focus on early, balanced bilinguals, participants had to meet the following criteria: they began learning both English and Mandarin before or by the age of 6 years old (Chee, Caplan, Soon, Sriram, Tan, Thiel & Weekes, 1999a; Chee, Tan & Thiel, 1999b; Fabbro, 1999; Paradis, 2004, 2009; Stafford, 2011); they reported literacy in both

Table 1. *Language background characteristics for the bilingual group*

	AoE	Speaking	Listening	Reading
English	3.34 (2.57) [1.96, 4.57]	8.61 (1.55) [7.84, 9.38]	9.15 (.98) [8.61, 9.61]	9.15 (1.06) [8.54, 9.69]
Mandarin	0(0)	9.15 (1.14) [8.53, 9.69]	9.53 (.77) [9.07, 9.84]	8.15 (1.55) [6.76, 9.23]

Note. Values reflect means (and standard deviations) and [95% confidence intervals]. AoE = age of exposure, defined as self-reported age (in years), of first exposure to the language; the AoE range for English was 0–6 years. Speaking, listening, and reading represent self-reported proficiency (maximum value 10) in these modalities, obtained from the LEAP-Q (Marian et al., 2007). For comparisons between English and Mandarin for the four variables, all $ps > .15$.

languages (Bialystok, 2001; Sanz, 2007); they reported at least weekly hearing/speaking and reading/writing of both languages; and they reported similar self-ratings of proficiency (see Table 1) in the two languages (Marian, Blumenfeld & Kaushanskaya, 2007).

Just as with the monolinguals in Morgan-Short et al. (2012b), the bilinguals were also screened for their experience learning Romance languages, such as Spanish or French, since Brocanto2 was designed to be Romance language-like in its (morpho)syntactic features (Morgan-Short et al., 2015; Morgan-Short et al., 2012a; Morgan-Short et al., 2012b; Morgan-Short et al., 2010); see below. Specifically, all participants had to meet the following criteria in order to be eligible to participate: no more than three years of classes total in any Romance language, with no more than one year at university, and no more than two weeks of immersion in a Romance language-speaking area, and none at all of either in the last two years. These criteria were selected to minimize unwanted confounds of exposure to (and thus potential transfer of) linguistic knowledge from other Romance languages (while at the same time avoiding excluding too many participants).

During the study, data from four of the 17 bilingual participants were excluded: two did not reach the low proficiency criterion (see below in Study Procedure), one scored below the pre-established IQ criterion (85 or above as measured by the Kaufman Brief Intelligence Test; KBIT; Kaufman & Kaufman, 2004), and one could not complete the final session of the study due to technical issues with the electroencephalogram (EEG) system. Thus, data from 13 early Mandarin–English bilinguals (nine female, all right-handed; Oldfield, 1971) were included in the analyses. For more detailed descriptive information on this final sample of bilinguals see Table 1.

The bilinguals were trained and tested under the exact same conditions as the previously-tested matched English monolingual group (Morgan-Short et al., 2012a; Morgan-Short et al., 2012b; Morgan-Short et al., 2010). Like the monolingual group, all bilingual participants had normal or corrected-to-normal vision and no reported neurological, learning, or neurodevelopmental disorders.

For additional descriptive details on age, IQ, years of education, number of foreign languages studied, and days between sessions for both groups, see Table 2. As can be seen in Table 2, the bilinguals and monolinguals differed only in the number of foreign languages they had studied, with the monolinguals reporting having studied significantly more foreign languages in a classroom setting than the bilinguals (also see Discussion), $t(1,27) = 3.00, p = .006, d = 1.15$.

Materials and procedure

The artificial language: Brocanto2

The target language used in this study was Brocanto2. This artificial language was designed by Morgan-Short and colleagues, following Brocanto (Friederici et al., 2002b), for the purpose of investigating additional language learning in a well-balanced and well-controlled experimental setting (for more details, see Morgan-Short et al., 2015; Morgan-Short et al., 2014; Morgan-Short et al., 2012a; Morgan-Short et al., 2012b; Morgan-Short et al., 2010). Brocanto2 is composed of a small lexicon and a limited number of grammar rules, allowing it to be learned to a high level of proficiency over the course of several days. Thus it allows for an in-depth longitudinal (within-subject) examination of the trajectory of language learning, from low to high proficiency.

Brocanto2 has been shown to be a reasonable model of natural language learning. First, this fully productive artificial language follows universal requirements of natural languages: all the grammatical features of Brocanto2 are found in natural languages, such as Supyire (spoken in Mali), which has subject–object–verb word order, grammatical gender agreement, and post-nominal adjectives and determiners (Carlson, 1994). Moreover, Brocanto2 is presented auditorily, and participants use the language in both comprehension and production tasks. Brocanto2 was based on another artificial language, Brocanto, developed and investigated by Friederici and colleagues (Friederici et al., 2002b; Opitz & Friederici, 2002). Importantly, both Brocanto and Brocanto2 have

Table 2. Descriptive information for bilingual and monolingual groups

	Bilinguals ($n = 13$)		Monolinguals ($n = 16$)	
	$M(SD)$	95% CI	$M(SD)$	95% CI
Age	21.69 (3.98)	[19.84, 23.76]	24.25 (4.34)	[22.31, 26.62]
Non-verbal IQ	114.38 (11.4)	[108.07, 120.07]	116.37(12.44)	[110.31, 122.25]
Years of Education	16.0 (4.16)	[14.07, 18.31]	16.25 (2.81)	[15.0, 17.75]
Number of foreign languages studied*	.92 (.64)	[.61, 1.31]	2.06 (1.23)	[1.50, 2.68]
Days between S1 and S2	1.84 (.98)	[1.31, 2.38]	1.62 (1.41)	[.93, 2.31]
Days between S2 and S3	2.31 (.94)	[1.84, 2.84]	2.12 (1.31)	[1.50, 2.81]

Note. Values reflect means (and standard deviations) and [95% confidence intervals]. Non-verbal IQ assessed with the non-verbal sub-section of the KBIT (Kaufman & Kaufman, 2004). S1 = session 1, S2 = session 2, S3 = session 3. The monolingual and bilingual groups did not differ for any variable other than the number of foreign languages studied ($p = .006$).

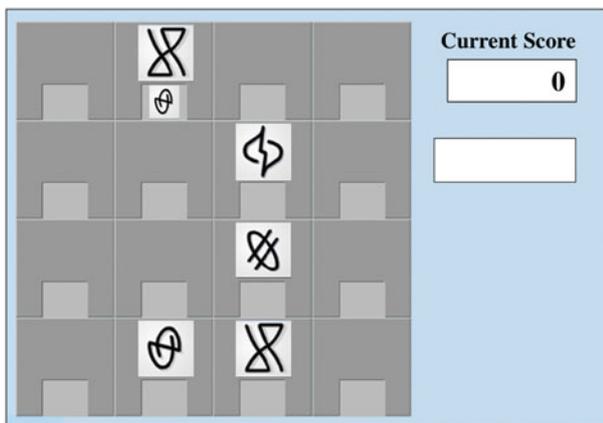


Figure 1. (Colour online) Example Brocanto2 game board and game piece configuration.

elicited neural patterns similar to those found in natural languages (Friederici et al., 2002b; Opitz & Friederici, 2002; Morgan-Short et al., 2015; Morgan-Short et al., 2012b; Morgan-Short et al., 2010).

Brocanto2 has a lexicon of 13 pronounceable words that can be combined to form 1,404 possible sentences, which refer to game moves in a computerized chess-like game (Morgan-Short et al., 2010; Morgan-Short et al., 2012b). See Figure 1 for a sample Brocanto2 game board. There are four nouns in Brocanto2, each of which refers to a game token - two are masculine (*pleck*, *neep*) and two are feminine (*vode*, *blom*; gender is not overtly marked on the nouns). There is one article, which is marked for linguistic gender (*li*, masculine; *lu*, feminine), and there are two adjectives, also marked for linguistic gender (*troise/neime* – masculine; *troiso/neimo* – feminine). The language also has two adverbs (*noyka*, *zayma*) as well as four verbs (*klin*, *praz*, *yab*, *nim*), which refer to the game moves (move, switch, release, capture, respectively) and differ with respect to transitivity.

Brocanto2 has fixed subject-object-verb word order at the sentence level and noun-(adjective)-article word at the noun-phrase level – making it a head-final language. A sentence such as *Pleck troise li vode troiso lu praz* means that “the round pleck switches with the round vode.” In the present study we focus on syntactic word order, which was tested in judgment tasks during the acquisition of EEG data (see below). As mentioned above, Brocanto2 was designed to be Romance language-like. Brocanto2 is similar to Romance languages (and different from both English and Mandarin) in its marking of grammatical gender on determiners and adjectives as well as in its post-nominal adjectives. In addition, the subject-object-verb word order in Brocanto2 is a possible word order in Romance languages (and different from both English and Mandarin), and at least one Romance language, i.e., Romanian, has post-nominal determiners. The fact that Brocanto2 has both post-nominal determiners and complete noun-phrase direct objects that occur pre-verbally seems to be unlike Romance languages (and also unlike English and Mandarin).

The lexicon of Brocanto2 follows English phonotactics, though none of the words exist in English. In Mandarin, *li* can either be a proper noun or a reference to a unit of distance (falling-rising tone means the distance is approximately 500 meters) and *lu* can have several meanings, depending on whether the tone is rising, falling, or rising-falling, and also on the sentence context. Since Mandarin is a tonal language (Zhou, Ye, Cheung & Chen, 2009), and meaning is encoded in the rising and falling tones on lexical items, the words *li* and *lu* as they are presented in Brocanto2 (always with even tone; never rising or falling) are unlikely to be confused with Mandarin words, and in fact no bilingual participants reported thinking that *li* or *lu* were Mandarin words.

Finally, note that participants did not see any written Brocanto2 stimuli during training, practice, or in the judgment task; that is, Brocanto2 was presented solely

auditorily, just as much of language learning and use is largely aural/oral.

Instructed training

Both bilinguals and monolinguals were trained in Brocanto2 under an Instructed condition. This was the exact same training condition that Morgan-Short and colleagues referred to as the ‘Explicit’ condition (Morgan-Short et al., 2012b; Morgan-Short et al., 2010). (We use the term ‘instructed’ rather than ‘explicit’ because the former seems to better describe the nature of the training context.) In the Instructed condition, which was designed to resemble a traditional foreign language classroom setting, participants received metalinguistic explanations about Brocanto2 grammar; for example, “In Brocanto2, both the subject and object are placed before the verb. The subject occurs first and the object occurs second.” Participants were also provided with 33 meaningful examples of the language: constellations or game moves presented together with the corresponding Brocanto2 sentence. This training lasted approximately 13 minutes.

Comprehension and production practice

Over the course of the study, participants engaged with 44 blocks of practice in Brocanto2: 22 blocks of comprehension practice and 22 of production. The practice blocks were completed in alternating blocks of two (two comprehension, then two production, then two comprehension, etc.). Each block consisted of 20 items, each of which corresponded to one game move. In comprehension practice, participants listened to a Brocanto2 sentence and had to make the corresponding game move on the computer screen using a mouse. In production practice, participants viewed a game move on the computer screen and were instructed to say aloud a Brocanto2 sentence that described that move. After each item, participants’ total score in the game increased or decreased by 10 points, depending on the accuracy of the comprehension or production response. The total game score was displayed on-screen during all practice sessions. Responses were scored by the researcher, who had an answer key to refer to, during production practice, and were scored automatically by the computer program during comprehension practice. During comprehension practice, response times were also acquired automatically for each item. Comprehension response time was measured as the time between the end of each Brocanto2 sentence and the time that the participant completed the corresponding game move.

Grammaticality judgment test

At two points in the study (once in Session 1 and again in Session 3, see Study Procedure just below), participants underwent behavioral/EEG assessments

Table 3. *Examples of grammatical and ungrammatical sentences in Brocanto2*

Grammatical	<i>Blom lu pleck li yab</i> Blom-piece the pleck-piece the release “The blom-piece releases the pleck-piece.”
Ungrammatical	<i>Blom lu *troiso li yab</i> Blom-piece the *round the release “The blom-piece releases the round.”

Note. * marks the point of ungrammatical word order.

in Brocanto2. Specifically, participants completed an untimed grammaticality judgment test (GJT) while EEG data were recorded. The GJT contained 240 sentences, 40 of which contained word order violations and 40 of which were matched correct (grammatical) control sentences. See Table 3 for examples of grammatical and ungrammatical Brocanto2 sentences.

The other 160 sentences presented gender agreement or verb argument violations, or were matched correct control sentences; these are not discussed here. Each word order violation sentence was created from its matched correct control sentence by replacing a word from one category (e.g., noun) with a word from a different category (e.g., adjective) that violated the word order structure of Brocanto2, and thus created an ungrammatical Brocanto2 sentence. The ungrammatical and grammatical (correct) word order sentences differed only with respect to this word order violation. There were four GJT lists, which were counter-balanced across participants and test sessions. For additional details on the design and balance of these items for the GJT, see Morgan-Short et al. (2012b). The following presentation sequence occurred for each sentence: A fixation cross appeared while a Brocanto2 sentence was presented through earphones. A prompt (“Good?”) was presented 500 ms after the last word of each sentence. Participants had up to five seconds to make a judgment with a mouse. The next sentence was presented immediately after their response.

Study procedure

After giving informed consent, participants underwent a screening session, during which information on language background and proficiency was collected, and memory and IQ were tested. One day to one month later, participants returned for Session 1 of the study, which began with pre-training. During pre-training they were taught the rules of the chess-like computer game, which provided a meaningful context in which Brocanto2 could be learned. The rules of the game are independent of the grammar rules of Brocanto2. (Playing the game consists of the comprehension and production-based game moves described above.) During pre-training, participants

read the rules of the game at their own pace and the computer program demonstrated the four possible game moves (capture, release, move, switch), which participants were then able to practice; game moves were made using a computer mouse. This pre-training was presented in English; no Brocanto2 or English-equivalent words were used. Immediately following the game pre-training, participants underwent naming pre-training. During naming pre-training participants were presented with each of the four game tokens (pleck, neep, vode, blom) on the screen and clicked on the token to hear its name. They were then given a naming test for all four tokens, which continued until they achieved 100% naming accuracy, as demonstrated by naming each token correctly three times in a row.

Immediately after pre-training, participants underwent Instructed training (described above). Following this, they completed comprehension and production practice until they reached above-chance accuracy on two consecutive comprehension blocks – referred to here as the point of “low proficiency” in Brocanto2. Once participants reached this criterion they completed the first GJT session during which EEG data were recorded.

Session 2 of the study took place one to five days after Session 1. This time period was selected to minimize forgetting while also maximizing participants’ scheduling availability given the time-intensive experimental design of the study (multiple sessions, each lasting several hours, over the course of several days). These time periods were statistically matched between the bilinguals and monolinguals (see Table 2). During Session 2, participants were again given Instructed training, which was identical to that provided in Session 1. After this training they completed comprehension and production practice up to block 36 of the total of 44 practice blocks.

Finally, Session 3 took place one to five days after Session 2. During this session, participants finished the last 8 blocks of practice (4 comprehension and 4 production blocks). Upon finishing these 8 practice blocks, participants completed the second GJT/EEG assessment. For ease of exposition this is referred to as the “high proficiency” point.

EEG data acquisition and analysis

EEG acquisition followed the same parameters as those used by Morgan-Short and colleagues (Morgan-Short et al., 2012a; Morgan-Short et al., 2012b; Morgan-Short et al., 2010). Scalp EEG was recorded in DC mode at a sampling rate of 500 Hz from 64 tin electrodes (10-20 system; Jasper, 1958), mounted in an elastic cap (Electro-Cap International, Inc., Eaton, OH). EEG was amplified using Neuroscan SynAmps and filtered online with a band-pass filter (DC to 100 Hz, 24-dB/octave attenuation), and offline using a 0.16-30 Hz band-pass filter. Scalp

electrodes were referenced online to the left mastoid, and impedances were kept below 5k Ω . Electrooculogram (EOG) activity was measured using free electrodes placed above and below the right eye (vertical EOG) and on the right and left canthi (horizontal EOG). Participants’ data from target words free of artifacts greater than 30 μ V in the EOG and greater than 75 μ V in the EEG were included in analyses.

ERPs were time-locked to the onset of the critical word for each sentence: i.e., the word that created the word order violation in ungrammatical sentences and the matched correct word in the grammatical sentences (e.g., blom lu *troiso / blom lu pleck; critical word underlined). ERPs were averaged offline in each participant for lateral and midline electrode sites using a 200 ms pre-stimulus baseline. Data processing was done using the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) plug-ins for MATLAB; statistical analysis was carried out in SPSS (version 22; IBM Corp.). We grouped lateral scalp electrodes into distributional regions (for similar approaches, see e.g., Chow & Phillips, 2013; Hahne & Friederici, 2001; Frenzel, Schlesewsky & Bornkessel-Schlesewsky, 2011; Tanner & Van Hell, 2014): on left scalp sites, left frontal1 (FP1, AF3, AF7), frontal2 (F3, F5, F7), frontal3 (FC3, FC5, FC7), central1 (C3, C5, T7), central2 (CP3, CP5, TP7), posterior1 (P3, P5, P7), posterior2 (PO3, O1, PO7); on right scalp sites, right frontal1 (FP2, AF4, AF8), frontal2 (F4, F6, F8), frontal3 (FC4, FC6, FT8), central1 (C4, C6, T8), central2 (CP4, CP6, TP8), posterior1 (P3, P5, P7), and posterior2 (PO4, O2, PO8). Data from midline sites comprised Fz, Cz, and Pz. Analyses were conducted separately on lateral and midline sites.

We analyzed three time-windows: 200–400 ms, 400–700 ms, and 700–1000 ms. The time-windows were selected on the basis of both visual inspection of the grand mean ERP waveforms and previous L1/L2 ERP research. They were similar to those used in previous Brocanto2 studies (e.g., Morgan-Short et al., 2010, Morgan-Short et al., 2012a; Morgan-Short et al., 2012b), and are generally representative of the ERP effects of interest (e.g., possible LANs, and P600s; for similar windows, see e.g., De Vincenzi, Job, Di Matteo, Angrilli, Penolazzi, Ciccarelli & Vespignani, 2003; Martin-Loeches, Muñoz, Casado, Melcon & Fernández-Frías, 2005; Van Den Brink & Hagoort, 2004).

Analyses across the two groups were conducted separately at each level of proficiency: low and high. This was done in order to clearly characterize bilingual vs. monolingual patterns at low and high proficiency while also keeping the ANOVAs within reasonable levels of complexity and power (see Luck & Gaspelin, 2017 for a discussion on overly-complex ANOVAs in ERP research). For lateral sites, the ANOVAs at each level of proficiency (low and high) included the within-subjects

Table 4. Summary of practice performance in bilinguals and monolinguals

	Bilinguals (<i>n</i> = 13)		Monolinguals (<i>n</i> = 16)	
	<i>M</i> (<i>SD</i>)	95% CI	<i>M</i> (<i>SD</i>)	95% CI
Block				
Blocks to LP	3.84(3.07)	[1.98, 5.71]	6.93(8.29)	[2.52, 11.35]
Blocks to 95%	7.9(5.9)	[4.3, 11.5]	9.3(8.2)	[4.9, 13.6]
Accuracy				
LP	59.44(16.02)	[49.75, 69.12]	55.34(23.01)	[43.09, 67.60]
Comp	86.73(10.60)	[80.32, 93.14]	82.83(16.67)	[73.95, 91.71]
Prod	90.65(6.07)	[86.98, 94.32]	82.32(19.93)	[71.69, 92.94]
Response time				
At low	15.63 (6.92)	[11.81, 19.54]	14.37 (4.63)	[12.07, 16.57]
At high	2.64 (1.48)	[1.83, 3.48]	2.91 (2.05)	[2.05, 3.98]

Notes. Values reflect means (and standard deviations) and [95% confidence intervals]. Accuracy represented as percent correct. Blocks to LP = number of production and comprehension blocks to reach low proficiency, 45% twice in a row (i.e., above chance performance); Blocks to 95% = number of production and comprehension blocks between low proficiency and reaching 95% twice in a row on practice; LP = accuracy at low proficiency; Comp = accuracy over all comprehension blocks; Prod = accuracy over all production blocks; At low = response time (in seconds) at low proficiency; At high = response time (in seconds) at high proficiency, that is, the last 4 comprehension blocks of practice, just prior to the high proficiency judgment/ERP assessment.

factors of Grammaticality (grammatical, ungrammatical), Hemisphere (left, right), and Anterior/posterior (frontal1, frontal2, frontal3, central1, central2, posterior1, posterior2), and the between-subjects factor Group (bilingual, monolingual). For the midline sites, ANOVAs at each level of proficiency included Grammaticality and Electrode (Fz, Cz, Pz) as within-subjects factors, and Group as the between subjects-factor. For both lateral and midline analyses, significant interactions with Grammaticality were followed up with step-down ANOVAs. We report Greenhouse-Geisser corrected *p*-values for data with more than one degree of freedom in the numerator. We report follow-up analyses on midline ANOVAs only if they provide unique information not revealed in the lateral analyses.

Behavioral data analysis

The behavioral data were analyzed following Morgan-Short and colleagues (Morgan-Short et al., 2012a; Morgan-Short et al., 2012b; Morgan-Short et al., 2010). For comprehension and production performance, accuracy (percent correct) data for comprehension and production, and reaction time data for comprehension, were analyzed. For the GJT, *d*-prime scores were calculated, providing an unbiased measure of participants' ability to discriminate between grammatical and ungrammatical items (Macmillan & Creelman, 2005; Stanislaw & Todorov, 1999; Wickens, 2002). These scores were calculated using the formula $d\text{-prime} = z(\text{hit rate}) - z(\text{false alarm rate})$. Data were analyzed using SPSS (version 22, IBM Corp.). Note that a *d*-prime

of zero indicates chance performance and a *d*-prime of four indicates near-perfect discrimination ability between grammatical and ungrammatical sentences.

Results

Practice performance

Accuracy: comprehension and production

Practice performance for the bilinguals and monolinguals is summarized in Table 4, and accuracy across all the practice blocks is depicted in Figure 2 (comprehension) and Figure 3 (production). As can be seen, by the end of practice both groups reached very high levels of accuracy in comprehending and producing Brocanto2. Descriptively, the bilinguals appear to reach low proficiency (above chance performance; see Methods) faster (i.e., after less practice) than monolinguals (blocks to LP, Table 4). Additionally, over the course of practice, the bilinguals appear to be more accurate than the monolinguals in both comprehension and production, at least after the first few blocks and until around blocks 17–18 (out of 22 in each modality; Figures 2 and 3), at which point the monolinguals catch up to the bilinguals' performance. Despite these descriptive differences suggesting superior performance by the bilinguals, ANOVAs comparing bilinguals and monolinguals on the block and accuracy variables summarized in Table 4, and planned *t*-tests comparing bilinguals and monolinguals on accuracy for each of the 22 comprehension and 22 production blocks, showed no statistical differences between the two groups, with all *ps*

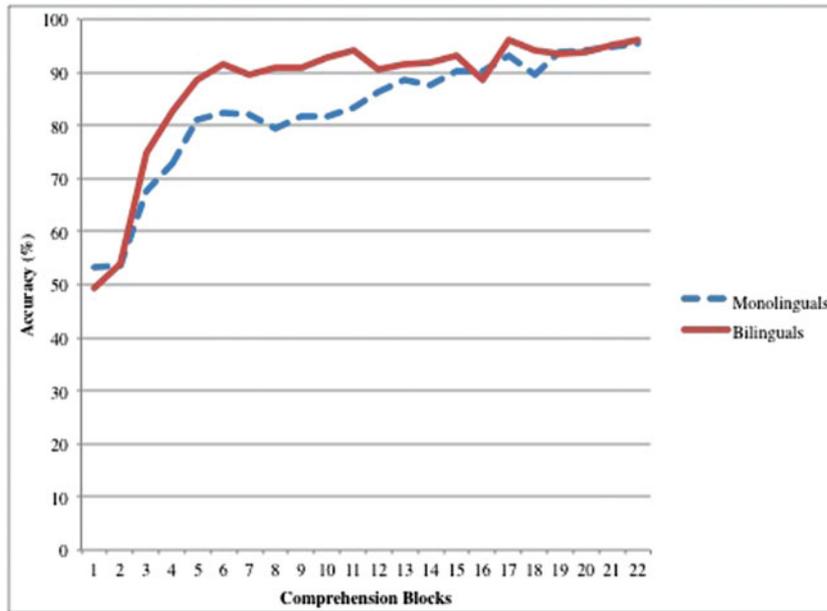


Figure 2. (Colour online) Comprehension accuracy for bilinguals and monolinguals, shown across the 22 comprehension practice blocks.

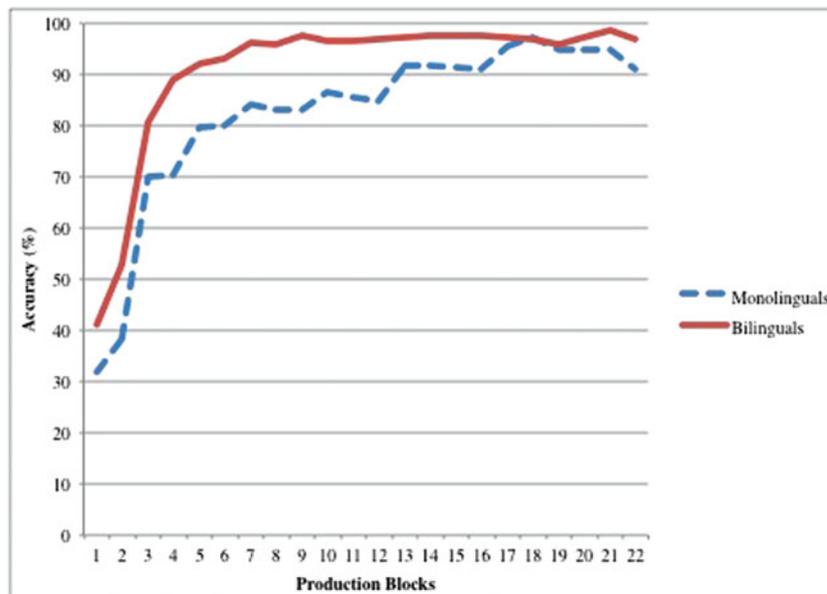


Figure 3. (Colour online) Production accuracy for bilinguals and monolinguals, shown across the 22 production practice blocks.

> .15 (for full descriptive statistics of each comprehension and production block see Appendix A).

Response time: comprehension

Although comprehension response time, defined as the time between the end of each Brocanto2 sentence and when the participant completed the corresponding game move (see Methods), is a somewhat coarse measure of

response speed, it is nevertheless useful for examining (a) whether participants responded more quickly over the course of learning and (b) whether the accuracy improvements outlined above might have occurred at a cost to processing time, i.e., whether there was a speed-accuracy tradeoff. Descriptively, similar to the results reported above for accuracy, the bilinguals generally responded faster than the monolinguals, at least after the first few blocks and until around block 18 (see

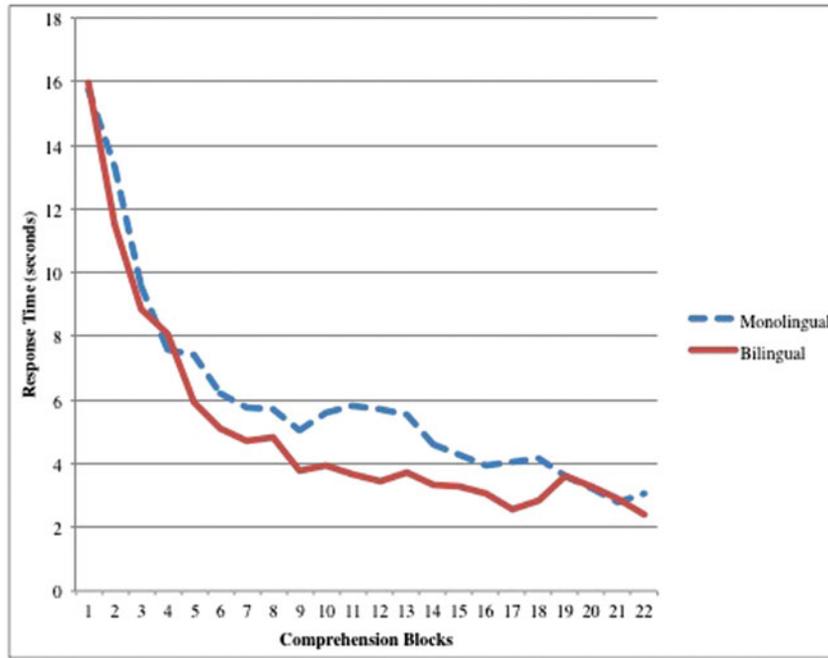


Figure 4. (Colour online) Comprehension response times for bilinguals and monolinguals, shown across the 22 comprehension practice blocks.

Figure 4). However, as with the accuracy data, ANOVAs comparing bilinguals and monolinguals on the response time variables summarized in Table 4, and planned *t*-tests comparing bilinguals and monolinguals on response time for each of the 22 comprehension blocks, showed no statistical differences between the two groups, with all *ps* > .15. (See Appendix A for full descriptive statistics of each comprehension block).

In sum, although descriptively the bilinguals seem to show better performance on comprehension and production and faster performance on comprehension, especially after the first few blocks and prior to around blocks 17–18, the statistical analyses did not reveal any bilingual/monolingual group differences in comprehension or production.

Grammaticality judgment test

Descriptive *d*-prime information for GJT performance at low proficiency (tested at the end of Session 1) and high proficiency (tested at the end of Session 3) is reported in Table 5. The mean *d*-prime scores are depicted in Figure 5. In separate analyses of *d*-prime scores at low and high proficiency (to parallel the separate ERP analyses at low and high proficiency) with Group (bilingual, monolingual) as the between-subjects factor, there were no significant differences between the groups at either time point (low proficiency, $F(1,27) = .771, p = .388, \eta_p^2 = .028$; high proficiency, $F(1,27) = .423, p = .521, \eta_p^2 = .015$). This was further confirmed by an ANOVA that included Time

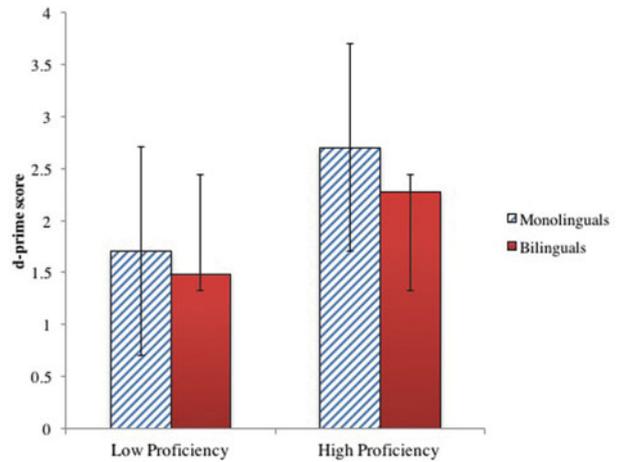


Figure 5. (Colour online) Mean grammaticality judgment performance at low and high proficiency in the bilingual and monolingual groups. Error bars show standard error.

(low, high) as a within-subjects factor and showed only a main effect of Time $F(1,27) = 22.62, p < .001, \eta_p^2 = .45$ with no other main effects or interactions (Time \times Group, $p = .877, \eta_p^2 = .001$; Group, $p = .425, \eta_p^2 = .024$).

Thus, as expected, there was an improvement in detecting word order violations from low to high proficiency. However, we did not observe bilingual/monolingual differences in grammaticality judgment at either low or high proficiency in Brocanto2. Note that although mean performance on the GJT may give the impression that monolinguals performed better,

Table 5. Summary of grammaticality judgment performance in bilinguals and monolinguals

	Bilinguals		Monolinguals	
	<i>M</i> (<i>SD</i>) <i>range</i>	95% CI	<i>M</i> (<i>SD</i>) <i>range</i>	95% CI
Low proficiency	1.25(1.49) -2.28-3.61	[.48, 2.03]	1.70(1.23) -.92-3.89	[1.01, 2.40]
High proficiency	2.24(1.92) -3.40-3.89	[1.08, 3.40]	2.62(1.23) .82-3.91	[1.97, 3.28]

Note. Values reported as *d*-prime. Values reflect means (standard deviations), ranges, and [95% confidence intervals].

the 95% confidence intervals for the GJT demonstrate that there is in fact considerable overlap for the two groups. Therefore, what appears to look like better monolingual performance on the behavioral measure is not a significant difference when examined at closer descriptive (95% CI) and statistical levels.

ERP results

At low proficiency

ERP waveforms and topographic voltage maps for word order processing at low proficiency in bilinguals and monolinguals are presented in Figure 6 and Figure 7. Visual inspection of the ERP waveforms suggested that ungrammatical word order (compared to matched grammatical word order sentences) elicited a P600 response in the bilingual group, with no corresponding ERP effect in the monolingual group (see Figures 6 and 7).

Results from the global ANOVAs for the three time-windows of interest², with Grammaticality, Hemisphere, and Anterior/posterior as within-subjects factors, and Group as the between-subjects factor, are presented in Table 6. Though there were no significant effects in the earlier time-windows (200-400 ms and 400-700 ms), there was a main effect of Grammaticality in the 700-1000 ms time-window. Visual inspection suggested this effect was being driven by the apparent P600 in the bilingual group.

To further examine this potential difference, and because low power in the sample might preclude detecting by-Group interactions, we conducted full follow-up ANOVAs within each group. In the bilinguals, the ANOVA for lateral sites showed a marginally significant Grammaticality \times Hemisphere interaction, $F(1,12) = 3.36$, $p = .091$, $\eta_p^2 = .22$ which was

² To address suggestions raised by an anonymous reviewer regarding possible negativities in the bilingual group at low and high proficiency, we also analyzed a 500-600 ms time-window at low proficiency in the bilingual group and a 300-600 ms time-window at high proficiency in the bilingual group. These analyses produced no significant main effects or interactions.

qualified by a significant Grammaticality \times Hemisphere \times Anterior/posterior interaction, $F(6,72) = 3.43$, $p = .046$, $\eta_p^2 = .22$. Follow-up analyses revealed a main effect of Grammaticality with significant (small-to-moderate) effects at central2 ($p = .030$, $\eta_p^2 = .34$), posterior1 ($p = .014$, $\eta_p^2 = .41$), and posterior2 ($p = .009$, $\eta_p^2 = .45$), confirming a P600 effect in the bilinguals. In contrast, in the monolinguals the follow-up ANOVAs produced no significant (or marginally significant) effects or interactions. In sum, the bilinguals but not the monolinguals showed a P600 in response to word order violations in Broca's area. (Note that the absence of the P600 in the monolinguals was also reported by Morgan-Short et al., 2012b with slightly different time windows.)

At high proficiency

Topographic voltage maps and ERP waveforms for word order processing at high proficiency in bilinguals and monolinguals are presented in Figures 7 and 8. Visual inspection of the waveforms and voltage maps suggested that word order violations elicited P600s in both groups. Additionally, visual examination suggested that the monolinguals showed an anterior positivity preceding the P600 effect whereas the bilinguals showed an anterior negativity.

Results from the ANOVAs at high proficiency in the three time-windows of interest are presented in Table 7. There were no significant effects in the 200-400 ms time-window. In the 400-700 ms time-window, there was a significant Grammaticality \times Group interaction. Follow-up analyses for this interaction confirmed an anterior positivity in the monolinguals: significant main effects of Grammaticality at frontal1 ($p = .011$, $\eta_p^2 = .36$) and frontal2 ($p = .011$, $\eta_p^2 = .36$), and a marginally significant effect at frontal3 ($p = .092$, $\eta_p^2 = .18$), consistent with the findings reported by Morgan-Short et al. (2012b). Follow-up analyses in the bilingual group yielded no significant results despite the visually-apparent anterior negativity.

In the 700-1000 ms time-window there was a main effect of Grammaticality and no significant interactions. This result suggested that ungrammatical word order

Table 6. *F*-statistics from the grand average ANOVAs on mean amplitudes at low proficiency in the 200–400 ms, 400–700 ms, and 700–1000 ms time windows

	<i>df</i>	200-400 ms	400-700 ms	700-1000 ms
<i>Midline</i>				
Gram.	1,27	–	–	3.934†
Gram. × Group	1,27	–	–	–
Gram. × Elec.	2,54	–	2.425†	–
Gram. × Elec. × Group	2,54	–	–	2.833†
<i>Lateral</i>				
Gram.	1,27	–	–	5.039*
Gram. × Group.	1,27	–	–	–
Gram. × Hemi.	1,27	–	–	3.272†
Gram. × Hemi. × Group	1,27	–	–	–
Gram. × Antpost	6,162	–	–	–
Gram. × Antpost × Group	6,162	–	–	–
Gram × Hemi × Antpost	6,162	–	–	–
Gram. × Hemi. × Antpost × Group	6,162	–	–	–

Note. Gram. = Grammaticality; Elec. = Electrode; Hemi. = Hemisphere; Antpost = Anterior/posterior; *df* = degrees of freedom. † $p < .10$; * $p < .05$.

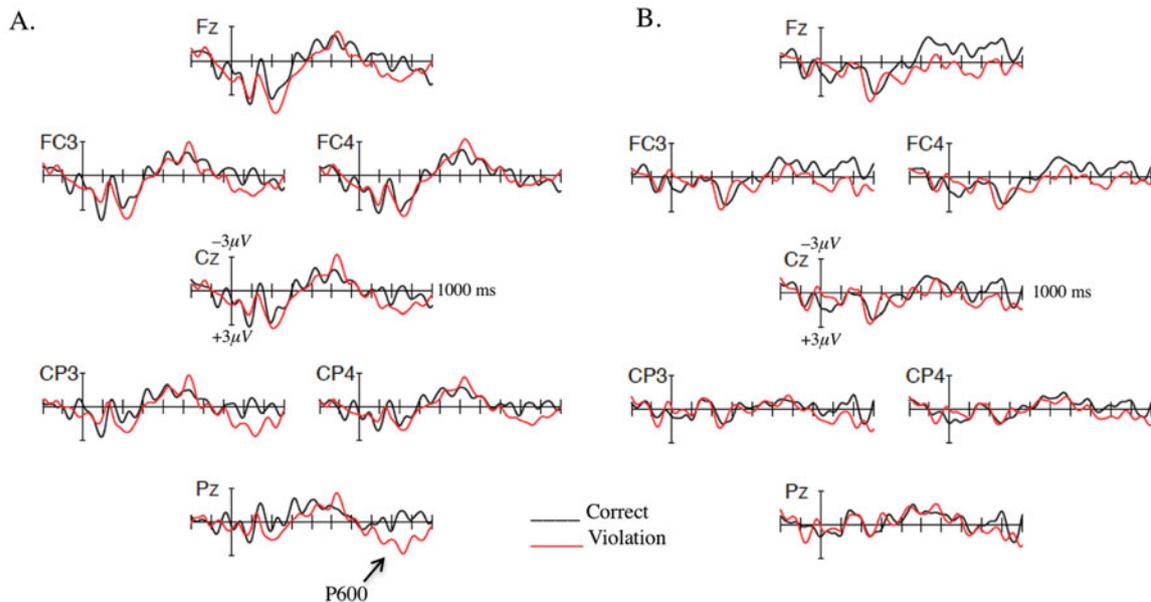


Figure 6. (Colour online) ERP waveforms at low proficiency. Panel A: Grand mean ERP waveforms in the bilingual group ($n = 13$) over 7 representative electrodes for correct word order (black line) and word order violation (red line) conditions at low proficiency. Panel B: Grand mean ERP waveforms in the monolingual group ($n = 16$) for correct word order (black line) and word order violation (red line) conditions at low proficiency. Each tick mark represents 100 ms. Note that negative voltage is plotted up and voltage is in microvolts (scale = $-3\mu V$ to $+3\mu V$).

elicited a P600 effect across both bilinguals and monolinguals, reflecting the visual interpretation of ERP effects in the two groups (see Figures 7 and 8). To examine this directly, we conducted follow-up ANOVAs within each group. The results of these follow-ups indeed showed a marginally significant P600 in the bilingual group,

$F(6,72) = 3.21, p = .088, \eta_p^2 = .21$ and a significant P600 in the monolingual group, $F(1,15) = 6.01, p = .027, \eta_p^2 = .29$, with similar effect sizes in the two groups.

In sum, at high proficiency in Broca2 both the bilinguals and monolinguals showed evidence of P600s in response to ungrammatical word order. However, the

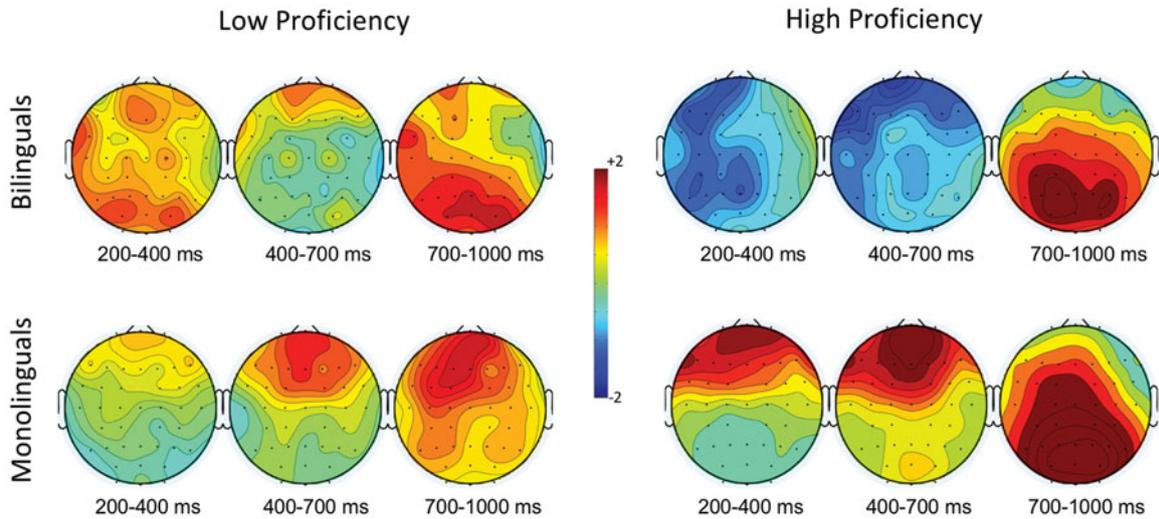


Figure 7. (Colour online) Topographic voltage maps. Maps show the scalp distribution of activity in the word order violation minus correct conditions, averaged for the 200-400 ms, 400-700 ms, and 700-1000 ms time-windows in each group at low and high proficiency. Calibration scale is $\pm 2\mu\text{V}$.

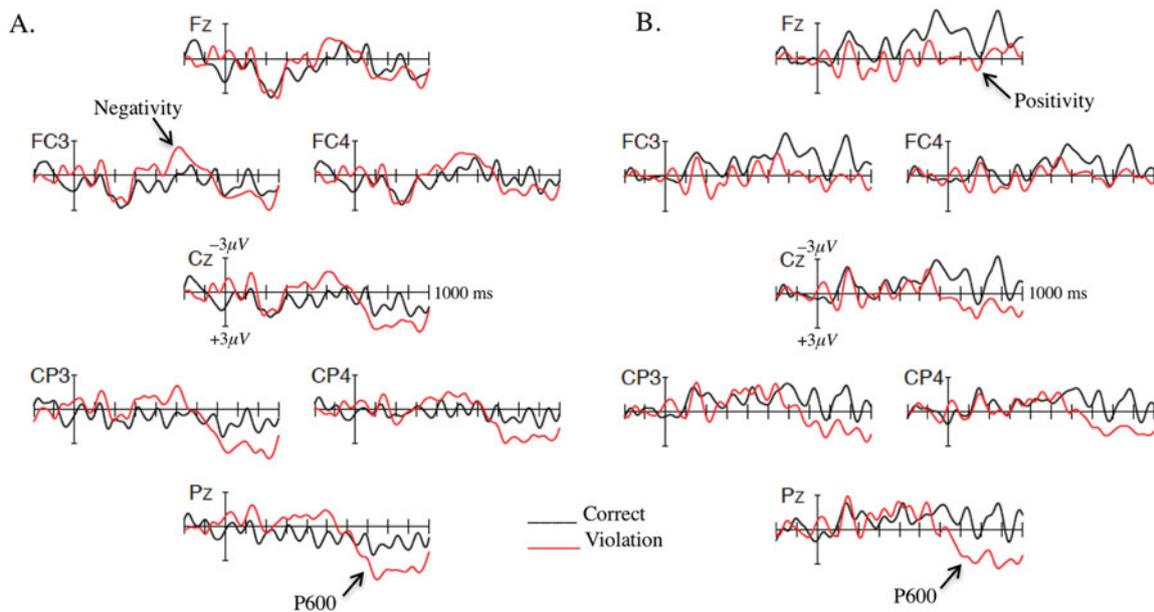


Figure 8. (Colour online) ERP waveforms at high proficiency. Panel A: Grand mean ERP waveforms in the bilingual group ($n = 13$) over 7 representative electrodes for correct word order (black line) and word order violation (red line) conditions at high proficiency. Panel B: Grand mean ERP waveforms in the monolingual group ($n = 16$) for correct word order (black line) and word order violation (red line) conditions at high proficiency. Each tick mark represents 100 ms. Note that negative voltage is plotted up and voltage is in microvolts (scale = $-3\mu\text{V}$ to $+3\mu\text{V}$).

groups differed in their ERP responses preceding the P600 effect, with monolinguals showing an anterior positivity that was not observed in the bilinguals.

Discussion

This study investigated potential behavioral and neural bilingual/monolingual differences in adult additional lan-

guage learning. There has been remarkably little research on potential differences in bilingual and monolingual additional language learning of grammar. Moreover, none of it has examined – as we did here – the trajectory of the learning or its neural correlates. Early Mandarin–English bilinguals and English monolinguals learned a well-studied artificial language, Brocanto2, and were compared on behavioral measures of comprehension, production,

Table 7. *F*-statistics from the grand average ANOVAs on mean amplitudes at high proficiency in the 200–400 ms, 400–700 ms, and 700–1000 ms time windows

	<i>df</i>	200-400 ms	400-700 ms	700-1000 ms
<i>Midline</i>				
Gram.	1,27	–	–	6.930*
Gram. × Group	1,27	–	3.104†	–
Gram. × Elec.	2,54	2.703†	–	4.187*
Gram. × Elec. × Group	2,54	–	–	–
<i>Lateral</i>				
Gram.	1,27	–	–	5.000*
Gram. × Group.	1,27	–	4.198*	–
Gram. × Hemi.	1,27	–	–	–
Gram. × Hemi. × Group	1,27	–	–	–
Gram. × Antpost	6,162	–	–	3.744†
Gram. × Antpost × Group	6,162	–	–	–
Gram × Hemi × Antpost	6,162	–	–	–
Gram. × Hemi. × Antpost × Group	6,162	–	–	–

Note. Gram. = Grammaticality; Elec. = Electrode; Hemi. = Hemisphere; Antpost = Anterior/posterior; *df* = degrees of freedom. † $p < .10$; * $p < .05$

and grammaticality judgment, as well as on their ERP patterns during syntactic processing at both low and high proficiency in the language. We first discuss the results for behavioral measures, followed by discussion of the ERP results.

Discussion of behavioral outcomes

The behavioral results revealed the following. The descriptive statistics for performance during practice suggested apparent bilingual advantages in language comprehension and production. However, statistical analyses did not reveal any significant differences in practice measures between the bilinguals and monolinguals. Similarly, though both groups improved in grammatical sensitivity to syntactic word order (as measured by grammaticality judgment tests) between low and high proficiency, there were no statistical differences between the groups on these tests. The finding that bilinguals and monolinguals did not differ significantly in behavioral measures of production and comprehension, or in grammaticality judgment of word order, contrasts with research that has shown bilingual advantages in language learning or performance in other domains, including vocabulary learning, pragmatic ability, and general proficiency (see Introduction; e.g., Antoniou et al., 2015; Safont-Jorda, 2003; Sanz, 2000).

The descriptive patterns in our behavioral measures nevertheless suggest the possibility of underlying group differences. In particular, the bilinguals appeared to show

superior language production and comprehension and faster comprehension than the monolinguals, especially after the initial practice blocks and before the last few blocks, when the monolinguals performed at similar levels as the bilinguals. This pattern partially fits with the small amount of research on bilingual L3 versus monolingual L2 learning of grammar. Cox (2017) found that older adult bilinguals outperformed matched monolinguals on sentence interpretation (e.g., comprehension), similar to the comprehension outcomes found here. Additionally, as in the current study, Cox (2017) found suggestive evidence of superior production, although that study assessed written sentence production whereas the current study assessed oral sentence production. Note also that Cox (2017) administered interpretation and written production assessments at discrete points (immediate and delayed posttests), in contrast to the continuous assessments we acquired here during comprehension and production practice. As indicated above, the difference we observed in bilingual and monolingual performance was most pronounced in the broad middle portion of learning (see Figures 2, 3, 4); by the end of practice both groups had essentially reached ceiling performance. Overall, then, the evidence suggests that there may be differences in the trajectory of learning for bilingual L3 as compared to monolingual L2 learning, with bilinguals reaching high levels of comprehension and production ability earlier than monolinguals, though further studies are needed.

The grammaticality judgment test yielded no significant differences between bilinguals and monolinguals in their discrimination between correct word order

and word order violations, at either low or high proficiency. These results align with prior work that has found no bilingual/monolingual differences in a grammaticality judgment test following artificial grammar learning (Nation & McLaughlin, 1986). Additionally, the results seem compatible with the findings reported by Cox (2017) of no bilingual/monolingual group differences in grammaticality judgment for participants that received explicit instruction. Nevertheless, at both low and high proficiency there was an apparent descriptive bilingual/monolingual group difference in the grammaticality judgment task. In particular, the monolinguals seemed to show better grammatical discrimination ability, on average, than bilinguals as measured by *d*-prime scores (see Table 5 and Figure 5). In the present study, recall that although the bilingual and monolingual groups were matched on a number of variables, there was a significant difference between monolinguals and bilinguals in the number of foreign languages they had studied, with the monolingual group having studied more languages in a classroom setting than the bilingual group (Table 2). One possible interpretation of the apparent GJT mean difference is that increased experience with classroom language learning, which was the exact setting that our Instructed training was designed to approximate, may have been advantageous to the monolingual group in performing the grammaticality judgments. However, even though the mean performance appears to be different, the 95% CIs of the two groups overlap substantially (see Table 5), underscoring the finding that the bilinguals and monolinguals in fact did not differ significantly in their grammatical judgment ability. Future studies may elucidate this issue.

Discussion of ERP outcomes

Despite the lack of clear behavioral differences in bilingual L3 and monolingual L2 learning, the ERP results revealed marked differences in bilinguals' and monolinguals' syntactic processing of the additional language. At low proficiency, the bilinguals showed a P600 in response to word order violations while the monolinguals did not. At high proficiency, although both groups showed P600s for word order violations, they diverged in their neural patterns preceding the P600s: in the monolinguals, word order violations elicited an anterior positivity, whereas the bilinguals did not exhibit this anterior positivity (and instead appeared to show an anterior negativity).

The P600 in the bilinguals at low proficiency is particularly striking when one considers that the monolingual group was not only matched to the bilingual group on various subject level factors (Table 2) but also on Brocanto2 practice performance at low proficiency (Table 4, LP). Moreover, the monolinguals were not

significantly different at judging the same word order violations. Given the general consensus of the P600 as an index of structural conflict reanalysis or resolution, the P600 observed in the bilingual group indicates that even at a very low (just above chance) level of linguistic proficiency, bilinguals are already employing neural mechanisms commonly associated with syntactic processing in native speakers of languages. As discussed in the Introduction, in monolingual L2 learning research, on both natural and artificial languages, P600s are not often found at lower L2 proficiency (for reviews see Bowden et al., 2013; Steinhauer et al., 2009). Indeed, several neurocognitive theories of L2 learning posit that higher levels of L2 proficiency are needed in order for L2 to approximate L1 patterns (e.g., Bates & MacWhinney, 1989; Green, 2003; MacWhinney, 2001; Ullman, 2005, 2015), but these do not discuss bilingual L3. The P600 observed at low proficiency in the bilinguals but not monolinguals indicates that perhaps bilingual L3 does not follow exactly the same trajectory as monolingual L2 learning that has been described in existing L2 research and related neurocognitive theories. Rather, even at quite low proficiency, the bilinguals' L3 neural processing may already resemble patterns observed in L1 speakers of languages, even when monolinguals do not (yet) show this pattern. Thus, the evidence suggests that when learning an additional language, bilinguals may more rapidly develop (as compared to monolinguals, for at least certain aspects of language) the neural processes found in native speakers of languages – even when bilinguals do not show significantly better performance. Future research that further examines the trajectory of bilingual L3 using fine-grained measures of processing, such as ERPs, should help to further elaborate these issues.

So why might the bilingual group, and not the matched monolinguals, show this neural sensitivity during syntactic processing in the additional language at low proficiency? A common explanation for observed behavioral advantages in bilingual compared to monolingual language learning posits that bilinguals benefit from greater metalinguistic awareness (Dillon, 2009; Jessner, 2008), beginning even in early childhood (e.g., Bialystok, 2001; Cummins, 1987). Although it is difficult to directly tie such an explanation to the P600 observed here in the bilingual group at low proficiency, one possibility is that the bilinguals possess greater metalinguistic awareness that aids them in linguistic processes that are reflected in the P600.

Another account for bilingual advantages in language learning proposes that bilinguals are more adept than monolinguals at managing cross-language interference while processing a newly-learned language. Bartolotti and Marian (2012) suggest that interference from other languages is one of the reasons that adult language learning seems to be so difficult. They

propose that bilinguals' naturalistic experience with two languages makes them better than monolinguals at controlling between-language competition during novel word learning. Even though they propose this explanation for bilingual advantages at word learning, it might also help explain the P600 observed here for syntactic processing in the bilinguals at low proficiency. In particular, the bilinguals' naturalistic experience with two languages may have made them more adept than the monolinguals at managing competition between their known syntax (both Mandarin and English for the bilinguals, or only English for the monolinguals) and the new Brocanto2 syntax. Better management of such cross-language syntactic competition may have manifested in the bilinguals' ability to engage reanalysis mechanisms early in language learning (i.e., at low proficiency).

At high proficiency, both bilinguals and monolinguals showed P600s during syntactic processing. The P600 in the monolingual group is consistent with prior L2 research showing that P600s are found more reliably at higher levels of L2 proficiency (Bowden et al., 2013; Steinhauer et al., 2009), and can be taken as evidence in support of neurocognitive theories of L2 that propose that increases in L2 proficiency enable processing to approximate patterns found for native languages (e.g., Bates & MacWhinney, 1989; Green, 2003; MacWhinney, 2001; Ullman, 2001, 2005). However, the monolinguals also showed an anterior positivity that was not observed in the bilingual group. This positivity has also been reported in other L2 ERP studies of (morpho)syntactic processing (Bowden et al., 2013; Mueller, Oberecker & Friederici, 2009), and may reflect increased engagement of attentional mechanisms (similar to the P3a component; Friedman, Cycowicz & Gaeta, 2001; Polich, 2007) in the monolingual group (Bowden et al., 2013; Morgan-Short et al., 2012b). The fact that the bilingual group did not exhibit this neural effect suggests that the bilinguals, unlike the monolinguals, did not need to engage additional extra-linguistic attentional mechanisms during syntactic processing. Note that research on non-linguistic bilingual/monolingual differences suggests that one dimension of a bilingual advantage in cognition is tied to superior attentional control in bilinguals compared to monolinguals (e.g., Bialystok, Craik, Klein & Viswanathan, 2004; Costa, Hernandez & Sebastian-Galles, 2008; but note Paap & Greenberg, 2013). This may help explain the anterior positivity in the monolinguals that was not observed in the bilinguals, but will need to be more directly tested in future research.

It is worthwhile to note that in the same time period during which the monolinguals showed the anterior positivity, the bilinguals appeared to show an anterior negativity (Figures 7 and 8). This visually apparent (but not statistically significant) negativity found for the bilinguals hints at the presence of a biphasic LAN-P600

response. As mentioned in the background, the biphasic LAN-P600 response is considered by some researchers to be a key neural signature of (morpho)syntactic processing in native speakers of languages (Friederici et al., 2002a; Molinaro et al., 2011). The potential presence of a LAN in the bilinguals, instead of the anterior positivity in the monolinguals, suggests that indeed very different processes are at play in the two groups. In particular, it seems to suggest that bilinguals may be engaging automatic syntactic processing mechanisms (reflected by the LAN) whereas the monolinguals appear to need to engage extra-linguistic attention-related processes (anterior positivity) to handle the same syntactic information. This possibility warrants further examination in future studies. Note also that ERP component overlap could in principle be masking a negativity in the monolinguals (due to a stronger positivity with the same spatiotemporal characteristics), or conversely a positivity in the bilinguals (overridden by a stronger negativity). Additionally, the apparent negativity in the bilinguals may have reduced the amplitude of their P600. For discussion of overlapping ERPs, see e.g., Roehm, Bornkessel-Schlesewsky, Rösler and Schlesewsky (2007) and Tanner and Van Hell (2014).

Directions for future research

As noted in the Introduction, very little research has examined bilingual L3 as compared to monolingual L2 learning of grammar, and none of this has probed the trajectory of learning or the underlying neural correlates, as was done in the current study. Thus, the initial study opens a variety of avenues for future research.

The present study explicitly controlled for several linguistic and experiential factors that would be interesting to independently investigate in future work. For example, English and Mandarin were selected to minimize cross-linguistic transfer issues to Brocanto2. However, future studies could manipulate cross-linguistic (dis-)similarity between the bilinguals' two languages and the target L3 of the study, in order to test for potential effects of this variable on bilingual L3 processing, and how it compares to monolingual L2 (for a review of L2 ERP studies and of cross-linguistic similarity, see Van Hell & Tokowicz, 2010).

The present study also controlled for the age of acquisition of the bilinguals' two languages, with only early bilinguals examined. Though our understanding of L3 learning and processing in early bilinguals deserves further study, investigation of the role of later (or more variable ages of) L2 acquisition on bilingual L3 learning and processing is also warranted. Bilingual proficiency in English and Mandarin was also controlled for, with only bilinguals who reported similarly high self-rated proficiency in reading, writing, speaking, and listening

for both languages being eligible. Future research could examine whether variation in the proficiency of the bilinguals' two languages might affect L3 learning and processing patterns.

Another potentially productive path for future research is whether different types of language training may influence bilingual L3 as compared to monolingual L2 grammar learning and processing. Whereas the present study employed a training condition that was designed to approximate a traditional foreign language classroom setting, it would be interesting to extend this work to other learning settings, such as immersion. Similarly, it seems important to extend the current study's investigation of bilingual L3/monolingual L2 learning and processing with controlled instructed input to actual foreign language classroom settings (i.e., outside of the laboratory context).

Finally, although the present study acquired brain measures longitudinally, these were gathered at discrete time points at low and high proficiency. It is also possible to acquire such measures continuously, throughout training (Batterink & Neville, 2013; Tagarelli, 2014), as was done here for behavioral measures of production and comprehension. This is an interesting possibility for future research in the examination of bilingual L3 and monolingual L2 learning.

Conclusion

To our knowledge, this is the first study to examine the neural (as well as behavioral) correlates of syntactic processing in bilingual L3 as compared to monolingual L2 learning, let alone to examine this behavioral and neural trajectory longitudinally. Although the descriptive pattern of the behavioral data suggested bilingual advantages at certain time points, the groups did not show any statistical differences in performance measures. In contrast, the ERP data revealed electrophysiological differences in bilingual and monolingual additional language learning of grammar, in at least two ways. First, even at a low level of proficiency in the additional language, bilinguals (and not the matched monolinguals) already demonstrated evidence of employing reanalysis mechanisms that are commonly observed during syntactic processing in native speakers of languages. Second, though both the bilinguals and monolinguals showed evidence of syntactic reanalysis at higher proficiency, the ERP patterns suggested that only the monolinguals also needed to engage extra-linguistic attentional mechanisms. Overall, the study supports the existing empirical trend suggesting bilingual advantages at additional language learning, and adds key insights to the still very limited existing research on grammar learning.

Appendix A

Table A1. *Descriptive statistics on accuracy (percent correct) for each of the 22 comprehension blocks*

	Bilinguals			Monolinguals		
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
1	49.23	22.98	[37.69, 61.92]	53.12	22.27	[43.13, 63.75]
2	53.84	23.84	[42.69, 66.53]	53.75	27.05	[40.64, 66.25]
3	75.0	23.0	[63.09, 86.53]	67.50	26.71	[54.70, 80.31]
4	82.69	19.21	[71.93, 91.92]	72.81	28.57	[57.81, 85.93]
5	88.46	14.77	[80.38, 95.76]	80.93	24.64	[68.44, 92.50]
6	91.53	10.87	[85.39, 96.91]	82.50	24.25	[70.31, 93.43]
7	89.61	12.98	[82.70, 95.76]	82.18	26.01	[68.75, 93.43]
8	90.76	14.26	[82.69, 97.69]	79.37	23.72	[67.51, 90.31]
9	90.76	14.97	[81.93, 97.69]	81.56	29.81	[65.63, 94.37]
10	92.69	10.53	[86.92, 98.07]	81.56	25.34	[68.75, 93.13]
11	92.23	8.62	[89.61, 98.07]	83.43	23.64	[70.94, 94.06]
12	90.38	10.09	[85.0, 95.76]	86.25	18.39	[76.56, 94.37]
13	91.53	12.31	[85.0, 97.31]	88.43	17.58	[78.43, 95.93]
14	91.92	11.82	[85.38, 97.31]	87.50	14.94	[79.69, 94.37]
15	93.07	11.09	[86.92, 98.07]	90.31	13.47	[82.82, 95.93]
16	88.46	10.48	[83.07, 93.84]	90.31	14.43	[82.19, 96.24]
17	96.15	6.81	[92.31, 99.23]	93.12	13.76	[85.31, 98.12]

Table A1. *Continued*

	Bilinguals			Monolinguals		
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
18	94.23	8.86	[89.23, 98.07]	89.68	11.02	[84.68, 94.37]
19	93.46	10.07	[87.69, 98.07]	93.75	11.61	[87.50, 98.43]
20	93.84	8.69	[88.84, 98.07]	94.06	10.2	[88.12, 98.12]
21	95.0	8.66	[90.38, 98.84]	94.68	7.84	[90.32, 97.81]
22	96.15	6.17	[92.69, 98.84]	95.31	8.45	[90.93, 98.75]

Note. Descriptive values reflect means, standard deviations, and 95% confidence intervals.

Table A2. *Descriptive statistics on response times for each of the 22 comprehension blocks*

	Bilinguals			Monolinguals		
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
1	15.93	8.34	[11.44, 20.28]	15.73	4.78	[13.44, 18.15]
2	11.49	6.78	[8.16, 14.93]	13.33	4.45	[11.32, 15.54]
3	8.86	6.43	[5.75, 12.49]	9.55	4.53	[7.53, 11.68]
4	8.06	7.11	[4.76, 12.62]	7.56	3.76	[5.85, 9.41]
5	5.94	4.58	[3.67, 8.54]	7.41	3.81	[5.75, 9.41]
6	5.08	3.57	[3.36, 7.01]	6.22	4.61	[4.32, 8.74]
7	4.71	3.21	[3.14, 6.41]	5.74	3.85	[4.12, 7.75]
8	4.81	3.75	[3.10, 6.75]	5.71	3.68	[4.16, 7.66]
9	3.79	2.31	[2.63, 4.99]	5.03	2.75	[3.82, 6.41]
10	3.94	2.39	[2.76, 5.17]	5.61	3.63	[4.06, 7.58]
11	3.66	2.04	[2.61, 4.76]	5.81	4.50	[3.89, 8.07]
12	3.43	1.43	[2.74, 4.13]	5.71	4.43	[3.84, 8.15]
13	3.70	2.22	[2.59, 4.95]	5.53	4.88	[3.60, 8.19]
14	3.31	1.92	[2.35, 4.38]	4.61	3.67	[3.25, 6.67]
15	3.26	1.77	[2.37, 4.25]	4.25	4.54	[2.64, 6.79]
16	3.04	1.65	[2.23, 3.88]	3.93	2.42	[2.86, 5.15]
17	2.56	1.56	[1.82, 3.52]	4.06	3.24	[2.70, 5.79]
18	2.84	1.58	[2.07, 3.71]	4.18	3.06	[2.88, 5.80]
19	3.62	2.28	[2.44, 4.92]	3.60	2.70	[2.52, 5.09]
20	3.26	2.09	[2.28, 4.35]	3.20	2.46	[2.20, 4.53]
21	2.90	1.92	[2.01, 4.06]	2.78	1.90	[1.98, 3.79]
22	2.37	1.26	[1.73, 3.11]	3.03	2.25	[2.07, 4.19]

Note. Response time represented in seconds. Descriptive values reflect means, standard deviations, and 95% confidence intervals.

Table A3. Descriptive statistics on accuracy (percent correct) for each of the 22 production blocks

	Bilinguals			Monolinguals		
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
1	41.15	34.41	[22.31, 59.23]	31.88	33.31	[16.56, 48.44]
2	52.69	36.66	[32.31, 70.37]	38.12	35.44	[20.63, 55.93]
3	88.85	17.45	[78.86, 96.92]	70.0	33.76	[52.50, 84.37]
4	88.46	14.77	[80.38, 95.38]	70.31	35.09	[51.56, 85.63]
5	91.92	11.28	[85.38, 96.92]	79.69	33.29	[60.95, 93.44]
6	93.08	11.64	[86.15, 98.08]	80.0	33.21	[61.25, 94.06]
7	96.15	4.16	[93.86, 98.08]	84.06	29.45	[67.51, 96.25]
8	95.76	7.31	[91.15, 99.23]	83.12	30.54	[65.63, 95.93]
9	97.69	3.30	[95.76, 99.23]	83.12	31.72	[65.31, 96.86]
10	96.54	4.73	[93.85, 98.85]	86.56	30.31	[69.38, 98.13]
11	96.54	4.73	[93.85, 98.46]	85.63	30.15	[68.44, 97.19]
12	96.92	5.96	[93.46, 99.62]	85.0	29.94	[68.13, 96.56]
13	97.31	5.63	[93.85, 99.62]	91.56	24.94	[77.51, 99.38]
14	97.69	3.30	[95.77, 99.23]	91.56	24.54	[78.44, 98.44]
15	97.69	3.30	[96.15, 99.23]	91.25	24.73	[77.51, 99.05]
16	96.69	3.88	[95.38, 99.62]	90.94	23.25	[78.75, 98.13]
17	97.31	2.59	[96.15, 98.85]	95.63	7.04	[91.88, 98.44]
18	96.92	4.80	[94.23, 99.23]	97.19	3.63	[95.31, 98.75]
19	95.77	6.07	[92.31, 98.46]	94.69	6.44	[91.25, 97.81]
20	97.31	5.99	[93.85, 100]	94.69	5.61	[91.88, 97.19]
21	98.46	2.40	[96.93, 99.62]	95.0	5.47	[93.19, 97.50]
22	96.92	4.35	[94.23, 99.23]	90.94	24.37	[78.44, 97.81]

Note. Descriptive values reflect means, standard deviations, and 95% confidence intervals.

References

- Abu-Rabia, S., & Sanitsky, E. (2010). Advantages of bilinguals over monolinguals in learning a third language. *Bilingual Research Journal*, 33(2), 173–199.
- Antoniou, M., Liang, E., Ettlinger, M., & Wong, P. C. M. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, 18(04), 683–695.
- Aparicio, X., Midgley, K. J., Holcomb, P. J., Pu, H., Lavaur, J.-M., & Grainger, J. (2012). Language effects in trilinguals: an ERP study. *Frontiers in psychology*, 3.
- Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: an ERP comparison. *Journal of Cognitive Neuroscience*, 17(1), 137–153.
- Bartolotti, J., & Marian, V. (2012). Language learning and control in monolinguals and bilinguals. *Cognitive Science*(Early view).
- Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in psychology*, 2.
- Bates, E. A., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). Cambridge, UK: Cambridge University Press.
- Batterink, L., & Neville, H. (2013). Implicit and explicit second language training recruit common neural mechanisms for syntactic processing. *Journal of Cognitive Neuroscience*, 25(6), 936–951.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy and cognition*. Cambridge: Cambridge University Press.
- Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, 19(2), 290–303.
- Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, 51(13), 2492–2511.
- Carlson, R. (1994). *A grammar of Supyire* (Vol. 14). Walter de Gruyter.
- Cenoz, J. (2013). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching*, 46, 71–86.
- Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism*, 7(1), 71–87.

- Cenoz, J., & Valencia, J. (1994). Additive trilingualism: Evidence from Basque Country. *Applied Psycholinguistics*, 15, 197–209.
- Chee, M. W., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W., Thiel, T., & Weekes, B. (1999). Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron*, 23(1), 127–137.
- Chee, M. W., Tan, E. W., & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *Journal of Neuroscience*, 19(8), 3050–3056.
- Chow, W.-Y., & Phillips, C. (2013). No semantic illusions in the “Semantic P600” phenomenon: ERP evidence from Mandarin Chinese. *Brain Research*, 1506, 76–93.
- Clahsen, H., & Felser, C. (2006a). Grammatical Processing in Language Learners. *Applied Psycholinguistics*, 27(1), 3–42.
- Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Science*, 10(12), 564–570.
- Costa, A., Hernandez, M., & Sebastian-Galles, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, 106(1), 59–86.
- Cox, J. G. (2017). Explicit instruction, bilingualism, and the older adult learner. *Studies in Second Language Acquisition*, 39(1), 29–58. doi: [10.1017/S0272263115000364](https://doi.org/10.1017/S0272263115000364)
- Cox, J. G., & Sanz, C. (2015). Deconstructing PI for the ages: Explicit instruction vs. practice in young and older adult bilinguals. *International Review of Applied Linguistics in Language Teaching*, 53(2), 225–248.
- Cummins, J. (1987). Bilingualism, language proficiency, and metalinguistic development. In P. Homel, M. Palif, & D. Aaronson (Eds.) *Childhood bilingualism: Aspects of linguistic, cognitive and social development*. Hillsdale, NJ; Erlbaum.
- De Angelis, G. (2007). *Third or Additional Language Acquisition*. Tonawanda, NY: Multilingual Matters.
- DeKeyser, R. M. (1996). Exploring automatization processes. *Tesol Quarterly*, 30(2), 349–357.
- DeKeyser, R. M. (1997). Beyond explicit rule learning. *Studies in Second Language Acquisition*, 19(02), 195–221.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1), 9–21.
- De Vincenzi, M., Job, R., Di Matteo, R., Angrilli, A., Penolazzi, B., Ciccarelli, L., & Vespignani, F. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and language*, 85(2), 280–296.
- Dillon, A. M. (2009). Metalinguistic awareness and evidence of cross-linguistic influence among bilingual learners in Irish primary schools. *Language Awareness*, 18(2), 182–197.
- Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. (2015). The relationship between artificial and second language learning. *Cognitive Science*. doi: [10.1111/cogs.12257](https://doi.org/10.1111/cogs.12257)
- Fabbro, F. (1999). *The neurolinguistics of bilingualism*. Hove, UK: Psychology Press.
- Falk, Y., & Bardel, C. (2010). The study of the role of the background languages in third language acquisition: The state of the art. *International Review of Applied Linguistics*, 48(2/3), 185–219.
- Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, 59, 226–248.
- Frenzel, S., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2011). Conflicts in language processing: A new perspective on the N400–P600 distinction. *Neuropsychologia*, 49(3), 574–579.
- Friederici, A., Hahne, A., & Saddy, D. (2002a). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, 31, 45–63.
- Friederici, A. D., & Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: First-pass and second-pass parsing processes. *Journal of Psycholinguistic Research*, 25(1), 157–176.
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002b). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, 99(1), 529–534.
- Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain’s evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, 25(4), 355–373.
- Gibson, M., Hufeisen, B., & Libben, G. (2001). Learners of German as an L3 and their production of German preposition verbs. In J. Cenoz Iragui, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition* (pp. 138–148). Clevedon: Multilingual Matters.
- Gillon Dowens, M., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish—Evidence from Event Related Potentials. *Neuropsychologia*, 49(7), 1651–1659.
- Gillon Dowens, M., Vergara, M., Barber, H., & Carreiras, M. (2009). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience*, 22(8), 1870–1887.
- Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology*, 33, 154–207.
- Gonzalez-Ardeo, J. (2000). Engineering students and ESP in the Basque Country: SLA versus TLA. In J. Cenoz Iragui, B. Hufeisen & U. Jessner (Eds.), *Looking beyond second language acquisition: Studies in tri- and multilingualism* (pp. 75–95). Tübingen: Stauffenberg.
- Green, D. W. (2003). Neural basis of the lexicon and the grammar in L2 acquisition. In R. v. Hout, A. Hulk, F. Kuiken & R. Towell (Eds.), *The interface between syntax and the lexicon in second language acquisition* (pp. 197–208). Amsterdam: John Benjamins.
- Grey, S., Tanner, D., & van Hell, J.G. (2017). How right is left? Handedness modulates neural responses during morphosyntactic processing. *Brain Research*, 1669, 27–43.

- Grey, S., Williams, J. N., & Rebuschat, P. (2014). Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition*(36), 611–645.
- Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental language learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences*, 38, 44–53.
- Guo, T., Ma, F., & Liu, F. (2013). An ERP study of inhibition of non-target languages in trilingual word production. *Brain and Language*, 127(1), 12–20.
- Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, 4, 123–141.
- Hama, M., & Leow, R. P. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition*, 32, 465–491.
- Jasper, H. H. (1958). The ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10, 371–375.
- Jessner, U. (2008). Teaching third languages: Findings, trends, and challenges. *Language Teaching*, 41, 15–56.
- Kaan, E. (2007). Event-related potentials and language processing: A brief overview. *Language and Linguistics Compass*, 1(6), 571–591.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test*: Wiley Online Library.
- Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition*(Early View).
- Kaushanskaya, M., & Marian, V. (2009a). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, 16(4), 705–710.
- Kaushanskaya, M., & Marian, V. (2009b). Bilingualism reduces native-language interference during novel-word learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, 35(3), 829–835.
- Kaushanskaya, M., & Rehtzigel, K. (2012). Concrete-ness effects in bilingual and monolingual word learning. *Psychonomic Bulletin & Review*, 19(5), 935–941.
- Kemp, C. (2007). Strategic processing in grammar learning: Do multilinguals use more strategies? *International Journal of Multilingualism*, 4, 241–261.
- Keshavarz, M. H., & Astaneh, H. (2004). The impact of bilinguality on the learning of English vocabulary as a foreign language. *International Journal of Bilingual Education and Bilingualism*, 7, 297–302.
- Lado, B. (2016). Aptitude and pedagogical conditions in the early development of a nonprimary language. *Applied Psycholinguistics* (early view), 1–23. doi: [10.1017/S0142716416000394](https://doi.org/10.1017/S0142716416000394).
- Lado, B., & Sanz, C. (2016). Methods in multilingualism research. In *Research Methods in Language and Education* (pp. 1–20). Springer International Publishing. doi: [10.1007/978-3-319-02329-8_14-1](https://doi.org/10.1007/978-3-319-02329-8_14-1)
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, 8.
- Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, 54, 146–157.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide, 2nd Edition*. New York, NY: Cambridge University Press.
- MacWhinney, B. (2001). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 69–90). Cambridge: Cambridge University Press.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and monolinguals. *Journal of Speech, Language, and Hearing Research*, 50(940-967).
- Martin, C., Strijkers, K., Santesteban, M., Escera, C., Hartsuiker, R., & Costa, A. (2013). The impact of early bilingualism on controlling a language learned late: an ERP study. *Frontiers in psychology*, 4.
- Martin-Loeches, M., Muñoz, F., Casado, P., Melcon, A., & Fernández-Frias, C. (2005). Are the anterior negativities to grammatical violations indexing working memory?. *Psychophysiology*, 42(5), 508–519.
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Reviews Neuroscience*, 7, 703–704.
- Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex*, 66(156-159).
- Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, 47(8), 908–930.
- Morgan-Short, K. (2014). Electrophysiological approaches to understanding second language acquisition: A field reaching its potential. *Annual Review of Applied Linguistics*, 34, 15–36.
- Morgan-Short, K., & Tanner, D. (2014). Event-related potentials (ERPs). *Research methods in second language psycholinguistics*, 127–152.
- Morgan-Short, K., Deng, Z., Brill-Schuetz, K. A., Faretta-Stutenberg, M., Wong, P., & Wong, F. C. K. (2015). A view of the neural representation of second language syntax through artificial language learning under implicit contexts of exposure. *Studies in Second Language Acquisition*, 37(02), 383–419.
- Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, 17(01), 56–72.
- Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. T. (2012a). Second Language Processing Shows Increased Native-Like Neural Responses after Months of No Exposure. *PLoS ONE*, 7(3), e32974.
- Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012b). Explicit and implicit second language training differentially affect the achievement of native-like brain

- activation patterns. *Journal of Cognitive Neuroscience*, 24(4), 933–947.
- Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, 60(1), 154–193.
- Mueller, J. L., Oberecker, R., & Friederici, A. D. (2009). Syntactic learning by mere exposure: An ERP study in adult learners. *BMC Neuroscience*, 10(89).
- Nation, R., & McLaughlin, B. (1986). Novices and Experts: An Information Processing Approach to the “Good Language Learner” Problem. *Applied Psycholinguistics*, 7(1), 41–55.
- Nayak, N., Hansen, N., Krueger, N., & McLaughlin, B. (1990). Language-learning strategies in monolingual and multilingual adults. *Language Learning*, 40(2), 221–244.
- Okita, Y., & Jun Hai, G. (2001). Learning of Japanese Kanji characters by bilingual and monolingual Chinese speakers. In J. Cenoz Iragui, B. Hufeisen & U. Jessner (Eds.), *Looking beyond second language acquisition: Studies in tri- and multilingualism* (pp. 63–73). Tübingen: Stauffenburg.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.
- Opitz, B., & Friederici, A. (2002). Artificial language acquisition: Changes in brain activity during the course of learning. *Journal of Cognitive Neuroscience, Supplement*, 35.
- Osterhout, L., Kim, A., & Kuperberg, G. R. (2012). The neurobiology of sentence comprehension. In M. Spivey, M. Joannisse & K. McCrae (Eds.), *The Cambridge handbook of psycholinguistics* (pp. 365–389). Cambridge: Cambridge University Press.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34(6), 739–773.
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232–258.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam, Netherlands: John Benjamins.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages* (Vol. 40): John Benjamins Publishing Company.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128–2148.
- Rah, A. (2010). Transfer in L3 sentence processing: Evidence from relative clause attachment ambiguities. *International Journal of Multilingualism*, 7, 147–161.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.
- Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, 81, 115–119.
- Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, 33(4), 829–856.
- Roehm, D., Bornkessel-Schlesewsky, I., Rösler, F., & Schlewsky, M. (2007). To predict or not to predict: Influences of task and strategy on the processing of semantic relations. *Journal of Cognitive Neuroscience*, 19(8), 1259–1274.
- Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 18(12), 2030–2048.
- Rossi, S., Hartmüller, T., Vignotto, M., & Obrig, H. (2013). Electrophysiological evidence for modulation of lexical processing after repetitive exposure to foreign phonotactic rules. *Brain and Language*, 127(3), 404–414.
- Rothman, J., Iverson, M., & Judy, T. (2011). Introduction: Some notes on the generative study of L3 acquisition. *Second Language Research*, 27(1), 5–19.
- Safont Jorda, M. P. (2003). Metapragmatic awareness and pragmatic production of third language learners of English: A focus on request acts realizations. *International Journal of Bilingualism*, 7(1), 43–68.
- Sanz, C. (2000). Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied Psycholinguistics*, 21, 23–44.
- Sanz, C. (2007). The role of bilingual literacy in the acquisition of a third language. In C. Perez-Vidal, A. Bel & M. J. Gauau (Eds.), *A Portrait of the Young in the New Multilingual Spain* (pp. 220–240). Clevedon, UK: Multilingual Matters.
- Sanz, C., & Lado, B. (2008). Third language acquisition research methods. In K. King & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd Edition ed., Vol. 10, pp. 3314–3336): Springer: US.
- Sanz, C., Park, H. I., & Lado, B. (2015). A functional approach to cross-linguistic influence in ab initio L3 acquisition. *Bilingualism: Language and Cognition*, 18(02), 236–251.
- Stafford, C. A. (2011). Bilingualism and enhanced attention in early adulthood. *International Journal of Bilingual Education and Bilingualism*, 14(1), 1–22.
- Stafford, C. A., Bowden, H. W., & Sanz, C. (2012). Optimizing language instruction: Matters of explicitness, practice, and cue learning. *Language Learning*, 62(3), 741–768.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavioral Research Methods, Instruments, & Computers*, 31(1), 137–149.
- Steinhauer, K. (2014). Event-related potentials (ERPs) in second language research: a brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L2. *Applied Linguistics*, 35, 393–417.
- Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, 120(2), 135–162.
- Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, 25(1), 13–41.
- Steinhauer, K., & Connolly, J. F. (2008). Event-related potentials in the study of language. In H. Whitaker (Ed.) *Concise Encyclopedia of Brain and Language* (pp. 91–104). Elsevier: Amsterdam.

- Tagarelli, K. M. (2014). *The neurocognition of adult second language learning: An fMRI study*. Unpublished doctoral dissertation. Georgetown University, Washington, DC.
- Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement. *Cortex*, *66*(149–155).
- Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in on-line L2 grammatical comprehension. *Bilingualism: Language and Cognition*, *17*, 277–293.
- Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, *16*(2), 367–382.
- Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289–301.
- Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *SSLA*, *27*(2), 173–204.
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*(1), 105–122.
- Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and Context in Adult Second Language Acquisition: Methods, Theory and Practice* (pp. 141–178). Washington, DC: Georgetown University Press.
- Ullman, M. T. (2015). The declarative/procedural model: A neurobiologically motivated theory of first and second language. In B. VanPatten and J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed.). New York: Routledge. pp. 135–158.
- van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010). Monitoring in language perception: mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, *22*(1), 67–82.
- Van Den Brink, D., & Hagoort, P. (2004). The influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs. *Journal of Cognitive Neuroscience*, *16*(6), 1068–1084.
- Van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: Syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, *26*(1), 43–74.
- Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Frontiers in psychology*, *5*.
- White, E. J., Genesee, F., & Steinhauer, K. (2012). Brain responses before and after intensive second language learning: Proficiency based changes in language background effects in adult learning. *PLoS ONE*, *7*(12), e52318.
- Wickens, T. (2002). *Elementary signal detection theory*. Oxford: Oxford University Press.
- Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, *27*, 269–304.
- Zare, M., & Mobarakeh, S. D. (2013). Effects of bilingualism on L3 vocabulary learning among Iranian EFL learners. *GEMA Online Journal of Language Studies*, *13*(1), 127–138.
- Zhou, X., Ye, Z., Cheung, H., & Chen, H.-C. (2009). Processing the Chinese language: An introduction. *Language and Cognitive Processes*, *24*(7–8), 929–946.