# Native language change during early stages of second language learning

Kinsey Bice and Judith F. Kroll



**Research on proficient bilinguals has demonstrated that both languages are always active, even when only one is required. The coactivation of the two languages creates both competition and convergence, facilitating the processing of cognate words, but slowing lexical access when there is a requirement to engage control mechanisms to select the target language. Critically, these consequences are evident in the native language (L1) as well as in the second language (L2). The present study questioned whether L1 changes can be detected at early stages of L2 learning and how they are modulated by L2 proficiency. Native English speakers learning Spanish performed an English (L1) lexical decision task that included cognates while event-related potentials were recorded. They also performed verbal fluency, working memory, and inhibitory control tasks. A group of matched monolinguals performed the same tasks in English only. The results revealed that intermediate learners demonstrate a reduced N400 for cognates compared with noncognates in English (L1), and an emerging effect is visually present in beginning learners as well; however, no behavioral cognate effect was present for either group. In addition, slower reaction times in English (L1) are related to a larger cognate N400 magnitude in English (L1) and Spanish (L2), and to better inhibitory control for learners but not for monolinguals. The results suggest that contrary to the claim that L2 affects L1 only when L2 speakers are highly proficient, L2 learning begins to impact L1 early in the development of the L2 skill.** ***NeuroReport*** **26:966–971** 

**NeuroReport** 2015, **26**:966–971

**Keywords: bilingualism, cognate facilitation, event-related potentials, inhibitory control, N400, second language learning**

Department of Psychology, Center for Language Science, The Pennsylvania State University, University Park, Pennsylvania, USA

Correspondence to Kinsey Bice, MS, Department of Psychology, Center for Language Science, The Pennsylvania State University, University Park, PA 16802, USA
Tel: +1 814 863 0126; fax: +1 814 863 7002; e-mail: klb489@psu.edu

**Received** 27 July 2015 **accepted** 14 August 2015

Proficient bilinguals are not like monolinguals of either language [1]. The parallel activation of their languages and the requirement to regulate the resulting competition change the native language (L1) by converging with the second language (L2) [2] and slowing or reducing lexical access [3]. These effects have also been demonstrated in late L2 learners who achieve high proficiency, but the assumption has been that a certain proficiency threshold must be reached before the effects of L2 on L1 can be observed. In support of this assumption, previous research on trilinguals reported that L1 cognates (words with orthographic and semantic overlap across languages; e.g. piano, or *clase* in Spanish) are facilitated by overlap with L2 when bilinguals are highly proficient and also by overlap with an L3, but only when there is high proficiency in the L3 [4].

Event-related potentials (ERPs) are a sensitive tool for investigating the earliest time course of processing. Importantly, they have been used to detect evidence of learning before behavioral measures [5]. Midgley *et al.* [6] used ERPs to investigate how L1 cognates are processed in intermediate L2 learners at a proficiency level for which previous research found no behavioral facilitation. These learners, however, revealed sensitivity to cognates during L1 processing in the form of a reduced N400 component. The N400 component is sensitive to the ease of lexical retrieval, where a reduced amplitude suggests facilitation and a larger amplitude suggests difficulty; cognates should be facilitated by the overlap across languages [7]. As previous research has demonstrated that L2 learning is evident in ERPs before behavior [5], this study showed that ERPs are also sensitive to changes in L1 processing during L2 learning earlier than previously assumed.

The present study aimed to extend this research to the earliest stages of L2 learning. As proficient bilinguals reveal changes in L1, we hypothesized that regulating the influence of L1 may be a critical step in L2 learning, and that individuals who adjust L1 processing earlier in learning may acquire higher L2 proficiency. We predicted facilitation in L1 in the form of a reduced N400 for L1 cognates, but not necessarily in reaction times (RTs) among beginning learners.

## Methods

### Participants

Sixty-one right-handed native English speakers (42 female, 20.47 ± 2.42 years) from the Pennsylvania State University participated in this study. Twenty-five (15 female, 20.4 ± 2.7 years) were functionally monolingual,


DOI: 10.1097/WNR.0000000000000453

**Table 1 Proficiency and cognitive characteristics of the participant groups**

| | Monolinguals | Beginning learners | Intermediate learners |
|---|---|---|---|
| Proficiency measures | | | |
| L1 self-rating score | 9.4 (0.6) | 9.4 (0.9) | 9.8 (0.5) |
| L1 verbal fluency score | 19.7 (2.8) | 19 (3.2) | 19.1 (4.4) |
| L2 self-rating score | 2.2 (1.7) | 4.7 (1.4) | 6.9 (1.3) |
| L2 verbal fluency score | NA | 4.6 (3.2) | 9.2 (2.9) |
| Cognitive measures | | | |
| O-Span score (correct recalled) | 44.5 (9.3) | 41 (14) | 48.5 (8) |
| Flanker effect (ms) | 41.3 (19) | 59 (48.3) | 44.7 (8) |
| Reactive inhibitory control (ms) | 205.3 (98.7) | 165.2 (94) | 195.2 (86.4) |

Values in parentheses are 1 SD. Self-rating scores range on a scale from 1 to 10. Verbal fluency scores are the average exemplars produced across five categories. The O-Span score is from the Operation Span task and has a maximum of 60. The flanker effect is the average incongruent response time minus the average congruent response time. The reactive inhibitory control measure is from the AX continuous performance task: average response time for AY trials (require reactive inhibitory control when Y, rather than X, appears) − average response time for BY trials (control condition).

and 36 were beginning ($n=23$; 14 female, $20.5\pm2.2$ years) and intermediate ($n=21$; 13 female, $20.6\pm2.5$ years) learners from Spanish classes. Neither group of learners differed from the monolinguals on cognitive measures (see Table 1). Learners were grouped using a median split of Spanish/English verbal fluency, which correlated with other subjective (e.g. self-rated L2 proficiency: $r=0.63$, $P<0.001$) and objective (e.g. $d'$ on the L2 lexical decision task: $r=0.72$, $P<0.001$) proficiency measures.

### Materials

The English and Spanish lexical decision task (LDT) contained 338 trials with 50% genuine words. Words (English mean length = 6.2, SD = 2.1; Spanish mean length = 6.07, SD = 1.63) and nonwords (English mean length = 6.5, SD = 2.1; Spanish mean length = 5.73, SD = 1.41) were matched in terms of length [English: $t$(311) = 1, $P=0.31$; Spanish: $t$(311) = 1.04, $P=0.3$] and onset, whenever possible. The words contained identical (e.g. piano; $N=19$) and nonidentical (e.g. crude, or *crudo* in Spanish; $N=19$) cognates and were matched with noncognate words in terms of length [English: $t$(70) = 0, $P=1$; Spanish: $t$(69) = 0.13, $P=0.9$] and frequency [English: $t$(70) = 0.16, $P=0.87$; Spanish: $t$(69) = 0.01, $P=0.99$]. Identical and nonidentical stimuli were combined during analyses for sufficient ERP power. Homographs and matched controls were included in the stimulus set but have not been reported here.

The category verbal fluency task in English and Spanish included five categories (fruits, vegetables, animals, body parts, clothing) in random order. Participants named as many items as possible in 1 min.

### Procedure

Participants completed the first session in English. Only learners returned for the second session in Spanish. Learners were told in the first session that no Spanish would be used, to reduce any anticipation of L2 use on L1 processing.

Session 1: participants completed the consent form and language history questionnaire. They performed the operation span task [8] that measured working memory, the Flanker task [9] and the AX continuous performance task ([10]) that measured inhibitory control and executive function, and the English verbal fluency task. These tasks were included to identify how individual differences in domain general functions relate to any observed change in L1.

Participants then performed the English LDT while EEGs were recorded from 32 electrodes using Neuroscan Quikcaps and software (Compumedics Neuroscan USA Ltd. Charlotte, North Carolina, USA), and a Synamps2 amplifier (Compumedics Neuroscan USA Ltd. Charlotte, North Carolina, USA) with a 24-bit analog to digital conversion (online sampling rate: 500 Hz; 0.05–100 Hz band-pass filter). The participant sat in a sound-attenuated, electrically shielded booth with a computer and button box. Researchers sat in an adjacent room with one computer to view EEG data and impedances and another to view stimulus presentation. Each trial began with a 500-ms fixation, 100-ms blank screen, followed by the presentation of the word in black letters on a white background, which disappeared upon a response or after 3 s.

Session 2: Spanish learners returned to perform the Spanish verbal fluency task and then the Spanish LDT, which used the same EEG parameters and trial procedures.

### Event-related potential data analysis

Continuous EEG data were analyzed using ERPLab [11]. Data were rereferenced offline to the average of both mastoids and subjected to a 30 Hz low-pass filter. Epochs from 200 ms before stimulus (baseline correction) until 800 ms after stimulus were extracted, and artifacts from eye or muscle movements were rejected. Rejected epochs (English: 14% of word trials, 11% of nonwords, 14% of noncognates, and 13% of cognates; Spanish: 8% words, 8% nonwords, 7% noncognates, 8% cognates) were double checked manually. Participants with fewer than 20 trials per condition were excluded.

Whole-head repeated-measures analyses of variance (ANOVAs) were conducted with group as a between-subject variable, electrode and word type as within-subject variables, and the mean amplitude from 300 to 500 ms as the dependent variable (DV). An ANOVA was carried out on the midline (Fz, Cz, Pz, Oz), first lateral (FP1, F3, FC3, C3, CP3, P3, O1, and right-hemisphere homologs), and second lateral (F7, FT7, T7, TP7, P7, and right-hemisphere homologs) electrodes. Only results that include a manipulated variable (group, word type) are reported, with a Greenhouse–Geisser correction for repeated-measures ANOVAs with more than two degrees of freedom in the numerator.

For correlational analyses, the mean amplitude of the difference wave across eight electrodes (CZ, PZ, C3, C4, CP3, CP4, P3, P4) from 300 to 500 ms was averaged to create a single representative value of the N400 magnitude for cognates compared with noncognates in each participant [12].

## Results

### English (L1) lexical decision task behavioral results

A one-way ANOVA with log-transformed RT as the DV revealed no group differences [$F(2,64)=0.63$, $P=0.54$]. However, an ANOVA with accuracy ($d'$) as the DV revealed a main effect of group [$F(2,64)=4.25$, $P=0.02$]. Bonferroni-corrected *t*-tests showed that intermediate learners (M = 5.57, SD = 1.57) were marginally more accurate than beginning learners (M = 4.52, SD = 1.59) and significantly more accurate than monolinguals (M = 4.29, SD = 1.29), but the latter groups did not differ.

A repeated-measures ANOVA with log-transformed RTs was used to investigate behavioral sensitivity to cognates. No main effects [group: $F(2,64)=0.61$, $P=0.55$; word type: $F(1,64)=0.06$, $P=0.81$] and no interaction [$F(2,64)=1.25$, $P=0.29$] were found. A repeated-measures ANOVA with accuracy (% correct) as the DV revealed a main effect of word type [$F(1,64)=11.9$, $P<0.001$], such that cognates (M = 0.96, SD = 0.05) elicited higher accuracy than noncognates (M = 0.95, SD = 0.05), but no effect of group [$F(2, 64)=0.4$, $P=0.67$] and interaction [$F(2,64)=0.13$, $P=0.88$] was found.

### Event-related potential results for English (L1) cognates

The midline ANOVA revealed a main effect of word type [$F(1,57)=4.62$, $P=0.04$], in which noncognates were more negative than cognates. The first lateral ANOVA revealed a marginally significant main effect of word type [$F(1,57)=2.72$, $P=0.1$]. Although these main effects did not statistically interact with group, an interaction is visually present in the waveforms and scalp topography (see Fig. 1).

### Spanish (L2) lexical decision task behavioral results

A *t*-test on log-transformed RTs revealed that beginning and intermediate learners did not differ in overall speed [$t(40)=1.52$, $P=0.14$]. However, a *t*-test on the $d'$ scores revealed that intermediate learners (M = 2.53, SD = 1.12) were more accurate than beginning learners [M = 1.55, SD = 0.66; $t(30)=3.42$, $P<0.01$].

A repeated-measures ANOVA with log-transformed RTs as the DV was used to investigate behavioral sensitivity to cognates. No main effect of group [$F(1,40)=2.13$, $P=0.15$] was found. A main effect of word type [$F(1,40)=10.25$, $P<0.01$] revealed that cognates (M = 6.64, SD = 0.2) elicited faster RTs than noncognates (M = 6.68, SD = 0.19). The interaction was not significant [$F(1,40)=0.89$, $P=0.35$]. A repeated-measures ANOVA with accuracy (% correct) as the DV revealed a main effect of group [$F(1,40)=11.83$, $P<0.001$], such that intermediate learners (M = 0.81, SD = 0.12) were more accurate than beginning learners (M = 0.67, SD = 0.18). A main effect of word type [$F(1,40)=21.99$, $P<0.01$] revealed that cognates (M = 0.78, SD = 0.19) elicited higher accuracy than noncognates (M = 0.68, SD = 0.14). The interaction was not significant [$F(1,40)=0.23$, $P=0.63$].

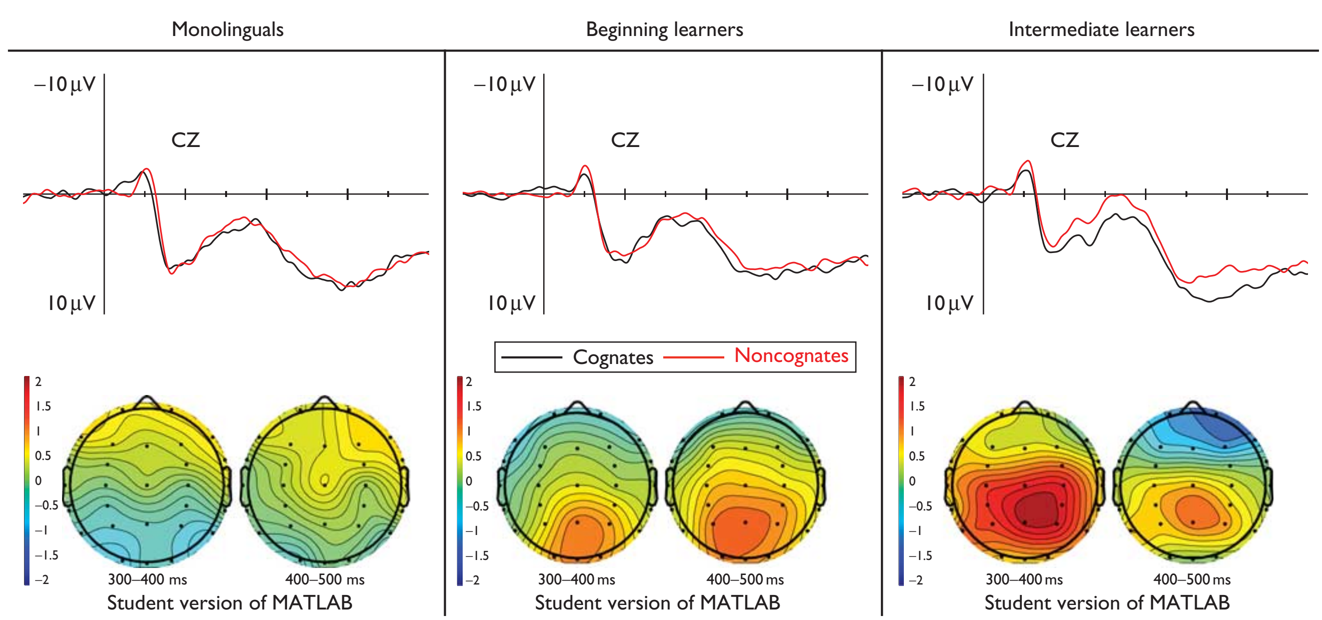
### Event-related potential results for Spanish (L2) cognates

The midline ANOVA revealed an interaction between word type and electrode [$F(3,105)=5.95$, $P<0.01$], in which noncognates were more negative than cognates across electrodes, particularly at frontal sites. A marginal interaction between group and word type emerged [$F(1,35)=3.33$, $P=0.08$], such that the N400 magnitude was larger in intermediate learners (M = 1.15) than in beginning learners (M = 0.08). The first lateral ANOVA revealed a main effect of word type [$F(1,35)=4.8$, $P=0.04$], with noncognates being more negative than cognates. The second lateral ANOVA revealed an interaction between word type and electrode [$F(4,140)=3.85$, $P=0.04$]. A visual inspection of the interaction showed that the magnitude of the N400 was the largest in frontal sites and disappeared in posterior sites. Waveforms and scalp topographies are shown in Fig. 2.

### Individual differences among learners

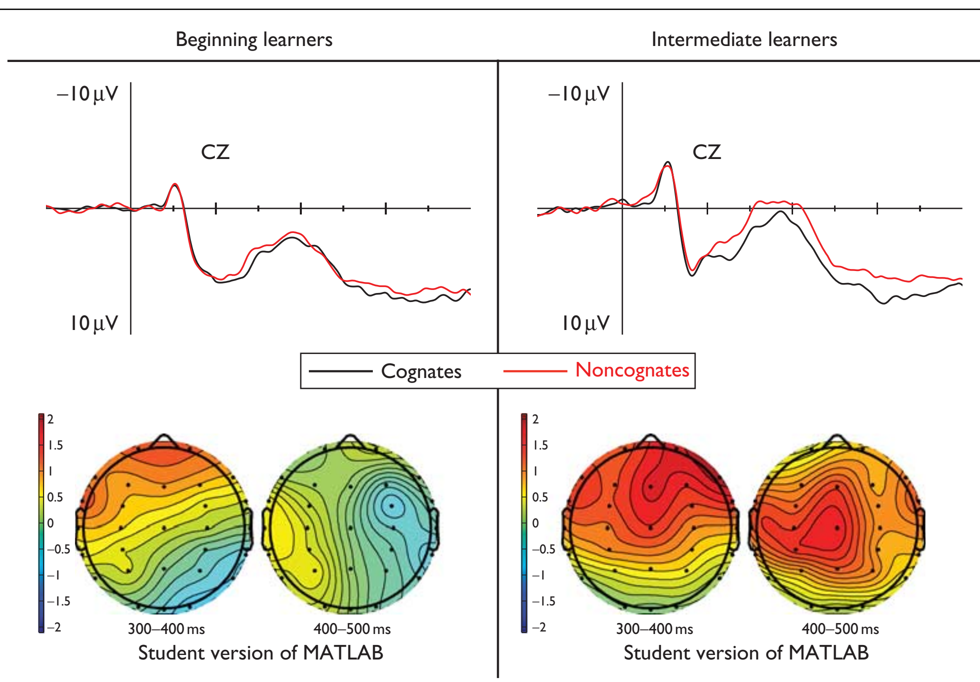
A regression model for the learners was built to understand the relationship between neural sensitivity to L1–L2 parallel activation, inhibitory control, and behavioral consequences of developing and managing the parallel activation during L2 learning (see Fig. 3). Log-transformed RTs from the English LDT were the outcome variables. The predictor variables were the magnitudes of the cognate N400 in English (L1) and in Spanish (L2), and the difference in RTs between AY trials and BY trials in the AX continuous performance task to index inhibitory control. The model [$F(3,29)=6.67$, $P<0.01$] showed that learners with larger cognate N400 magnitudes in English ($t=3.04$, $P<0.01$) and in Spanish ($t=3.18$, $P<0.01$) were slower to respond in English. Learners with better inhibitory control (i.e. a smaller RT difference between AY and BY trials) were also slower to respond in English (L1). There was no relationship between inhibitory control and RTs for monolingual participants ($r=-0.07$, $P=0.76$), or between English cognate N400 magnitudes and RTs ($r=-0.16$, $P=0.57$).

**Fig. 1**




ERPs for English (L1) cognates and noncognates. Top row: ERP waveforms from the representative electrode CZ for cognates (black line) and noncognates (red line) in English. Note that negative is plotted up. Bottom row: ERP scalp topography showing the distribution of the effect from 300 to 400 ms (left) and from 400 to 500 ms (right). Scale is from −2 (blue) to 2 (red) microvolts. Difference wave was calculated by subtracting noncognate amplitude from cognate amplitude; therefore, positivity (red) indicates that cognates have reduced negativity compared with noncognates. ERP, event-related potential.

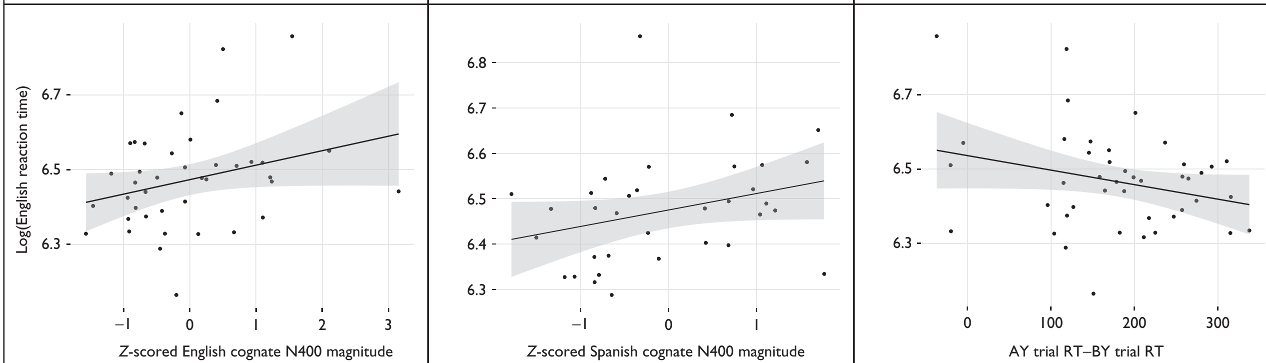
**Fig. 2**




ERPs for Spanish (L2) cognates and noncognates. Top row: ERP waveforms from the representative electrode CZ for cognates (black line) and noncognates (red line) in Spanish. Negative is plotted up. Bottom row: ERP scalp topography showing the distribution of the effect from 300 to 400 ms (left) and from 400 to 500 ms (right). Scale is from −2 (blue) to 2 (red) microvolts. Difference wave was calculated by subtracting noncognate amplitude from cognate amplitude; therefore, positivity (red) indicates that cognates have reduced negativity compared with noncognates. ERP, event-related potential.

**Fig. 3**

| English cognate N400 | Spanish cognate N400 | Reactive inhibitory control |
|---|---|---|




| $r=0.29$ $P=0.07$ | $r=0.3$ $P=0.08$ | $r=-0.27$ $P=0.08$ |
|---|---|---|

Model : $F(3,29)=6.67$, $P=0.001$
Regression line:
Log-transformed English RT = 6.48 + 0.05 (*Z*-scored English cognate N400 amplitude) + 0.06 (*Z*-scored Spanish cognate N400 amplitude) − 0.04 (*Z*-scored reactive inhibitory control)

Relationship between ERPs, inhibitory control, and behavioral costs. Correlation plots depicting the relationships captured in the regression model. The *y*-axis is the log-transformed reaction time from the English (L1) lexical decision task. The English and Spanish cognate N400 values (*z*-scored in the plots) are the mean amplitude from 300 to 500 ms of the difference wave of cognates minus noncognates, whereby more positive values indicate a larger cognate effect. The reactive inhibitory control value is created from the AX-CPT task: AY response time (reactive inhibitory control condition) − BY response time (control condition). *R*-values for the raw correlations are marginally significant, but when combined in the model, each becomes a significant predictor. The model statistics and slopes are provided at the bottom of the figure. ERP, event-related potential.

## Discussion

The present study used behavioral and ERP methods to investigate when and how L1 changes can first be observed as a result of L2 learning. The assumption has been that high L2 proficiency must be attained before bidirectional effects can be observed. However, the results we report challenge this assumption by demonstrating an emerging N400 for cognates in English (L1) in beginning learners of Spanish, which appears larger in intermediate learners. Although the N400 magnitude did not statistically interact with group, the waveforms and scalp distributions across groups capture the graded effect of L2 proficiency, and the N400 is absent in monolinguals. Furthermore, the magnitude of the English cognate N400 in learners is related to RTs, such that participants who were slower overall in English (L1) also showed a larger cognate N400 in English and had better inhibitory control. Critically, the relationship between RTs and inhibitory control is only present in learners, not in monolinguals, suggesting that experience with developing and managing cross-language competition during L2 learning engages inhibitory control in a way that monolingual language processing does not.

The primary finding of this study is that the earliest stages of L2 learning produce changes in the existing L1 network, as seen in the emerging cognate effect in beginning learners. The fact that this effect is not detected in behavior is not surprising; the L1 remains strongly dominant and any bidirectional effect should be subtle. The emerging N400 for beginning learners is small but present at the group level, and this is the first study to demonstrate this in L1 at such early stages of learning. Furthermore, ERPs for the intermediate learners nicely replicate the ERPs for L1 cognates in intermediate learners from a study by Midgley *et al.* [6], and the lack of behavioral effects are in line with a trilingual study by Van Hell and Dijkstra [4].

Another interesting finding is that, in Spanish (L2), beginning learners had a barely detectable N400 for cognates. The intermediate learners showed the expected effect, with cognates being significantly less negative than noncognates. This finding could be the result of very low proficiency in the beginning learners, such that noncognates are processed shallowly and lack semantics, reducing the amplitude of the noncognates. However, another possibility is that learners develop the skill to regulate the influence of L1 to benefit L2 learning, reducing the observed cognate effect in L2. Although these possibilities are not mutually exclusive, support for the latter explanation comes from the finding that the magnitudes of the cognate N400 in English and Spanish are negatively correlated. Learners with a canonical cognate N400 in English show a noncanonical N400 in Spanish (cognates more negative than noncognates). This suggests that some learners inhibit L1 during L2 use, making overlapping words across language more, rather than less, difficult to access. The role of inhibitory control in this account is supported by the regression model, in which better inhibitory control and greater parallel activation are related to slower English responses.

Finally, the discrepancy between neural and behavioral results is interesting. The effects on L1 were expected to be subtle. However, correlational analyses revealed no relationship between the neural (N400) and behavioral cognate effects in English. Inversely, beginning learners in Spanish did show a behavioral cognate effect, but not an N400. In Spanish, the neural and behavioral cognate effects were related, but in the unexpected direction: greater negativity for cognates compared with noncognates corresponded to cognate facilitation in RTs. The temporal, qualitative, and quantitative relationships between ERP and behavioral effects are still largely undetermined. Previous research on L2 has suggested that ERP effects precede behavior [5], providing a temporal relationship, but to what extent the magnitude of neural effects is related to the magnitude of behavioral effects, and the direction of that relationship require further investigation.

### Conclusion

The present study extends our knowledge of when and how the language network accommodates L2 learning. Previous research assumed a unidirectional (L1 to L2) influence in language processing until sufficient L2 proficiency was achieved, only after which bidirectional effects were thought to be observed. However, we used ERPs to detect subtle changes in L1 processing at much earlier stages of L2 learning. The development and incorporation of an L2 network into the established language network has subtle but detectable consequences for L1; namely, L1 becomes sensitive to the influence of L2, and processing slightly slows as a function of parallel activation and inhibitory control. Future research can track L1 changes within individuals over a period of L2 learning to capture the temporal dynamics of these effects and to identify individual differences in the trajectory of learning.

## Acknowledgements

The authors thank Gabrielle Abrams, Nicholas Shaff, and Emily Strauss for their assistance. This work and the writing of this paper were supported in part by NIH Grant HD053146 and NSF Grant OISE-0968369 to J.F. Kroll.

### Conflicts of interest

There are no conflicts of interest.

## References

1 Kroll JF, Dussias PE, Bice K, Perrotti L. Bilingualism, mind, and brain. *Annual Review of Linguistics.* 2015; **1**:377–394.
2 Ameel E, Malt BC, Storms G, Van Assche F. Semantic convergence in the bilingual lexicon. *J Mem Lang* 2009; **60**:270–290.
3 Kroll JF, Gollan H. Speech planning in two languages: what bilinguals tell us about language production. In: Ferreira V, Goldrick M, Miozzo M, editors. *The Oxford Handbook of Language Production* 2014. Oxford: Oxford University Press. pp. 165–181.
4 Van Hell JG, Dijkstra T. Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychon Bull Rev* 2002; **9**:780–789.
5 McLaughlin J, Osterhout L, Kim A. Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nat Neurosci* 2004; **7**:703–704.
6 Midgley KJ, Holcomb PJ, Grainger J. Effects of cognate status on word comprehension in second language learners: an ERP investigation. *J Cogn Neurosci* 2011; **23**:1634–1647.
7 Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu Rev Psychol* 2011; **62**:621–647.
8 Turner ML, Engle RW. Is working memory capacity task dependent? *J Mem Lang* 1989; **28**:127–154.
9 Eriksen CW. The flankers task and response competition: a useful tool for investigating a variety of cognitive problems. *Vis Cogn* 1995; **2**:101–118.
10 Braver TS. The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn Sci* 2012; **16**:106–113.
11 Lopez-Calderon J, Luck SJ. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci* 2014; **8**:213.
12 Tanner D, Van Hell JG. ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia* 2014; **56**:289–301.