Bilingualism: Language and Cognition

cambridge.org/bil

Desirable difficulties while learning collocations in a second language: Conditions that induce L1 interference improve learning

Manuel F. Pulido and Paola E. Dussias

Center for Language Science, Department of Spanish, Italian and Portuguese, Penn State University

Research Article

Cite this article: Pulido MF, Dussias PE (2020). Desirable difficulties while learning collocations in a second language: Conditions that induce L1 interference improve learning. *Bilingualism: Language and Cognition* **23**, 652–667. https://doi.org/10.1017/S1366728919000622

Received: 6 November 2018
Revised: 6 August 2019
Accepted: 24 August 2019
First published online: 11 October 2019

Key words:
multi-word units; second language learning; interference; lexical selection; L1 Regulation Hypothesis

Address for correspondence: Manuel F. Pulido, E-mail: mfp149@psu.edu

Abstract

Learning conventional verb-noun combinations in a second language is known to be highly problematic when word choices differ from those in the native language. Grounded on recent proposals of desirable difficulties in vocabulary learning (Bjork & Kroll, 2015), we tested Spanish learners of English on a new paradigm that aimed to induce interference from the native language during lexical selection in a second language, as a way to train regulation of the dominant language. Results showed that recall rates were significantly higher in the group of learners that practiced in conditions of L1-interference. Faster RTs showed more efficient lexical selection in those same learners. Additionally, RTs revealed that the more successful learners in both groups incurred a cost in accessing verb choices congruent with the native language, a finding that is consistent with an inhibitory account.

1. Introduction

An aspect of language that shows great variability in learning outcomes in a second language (L2) is that of collocations, defined as combinations of words that are associated due to frequent co-occurrence (Gries, 2013). For instance, the verb-noun (V-NP) sequence *run a business* is collocational in English, but *carry a business* is not. According to recent proposals (e.g., Arnon & Snider, 2010; Durrant & Doherty, 2010; Siyanova-Chanturia, Conklin, Caffarra, Kaan & van Heuven, 2017), the mental representation of collocations is different from that of single lexical items, but also from other fixed multi-word expressions, such as idioms (e.g., *kick the bucket*). Further, collocations have abstract internal syntax, and are said to influence language acquisition and language processing (Ibbotson, 2013).

Knowledge of L2 collocations is a critical asset for non-native speakers trying to master a foreign language. Learning them helps L2 speakers fulfill pragmatic functions and produce output that matches the expectations of native interlocutors (Bardovi-Harlig, 2009). Moreover, collocations are important because they facilitate L2 production (Boers, Eyckmans, Kappel, Stengers & Demecheleer, 2006) and processing (Wolter & Gyllstad, 2011). Although much is known about the special status that conventional speech has in the lexicon of native speakers, the cognitive mechanisms underlying how collocations come to be represented in adult L2 learners are not well understood. We know that L2 speakers are able to learn new single vocabulary items throughout life; they are, however, less successful at learning and using conventionalized collocations, even when collocations are composed of lexical items that, individually, are well known (Nguyen & Webb, 2017).

Despite the documented constraints associated with learning collocations in an L2 (e.g., Nesselhauf, 2003, 2005; Nguyen & Webb, 2017; Peters, 2016), some learners still manage to successfully achieve native-like performance in processing L2 collocations, but the evidence on what it takes for a learner to be successful, or for a learning context to be enabling, is mixed. Past work has shown that congruent collocations (i.e., those with the same lexical items in the L1 and L2) do not necessarily pose serious learning and processing challenges for L2 speakers; however, the development of L1-L2 incongruent representations is highly problematic (Nguyen & Webb, 2017; Peters, 2016). Even when learned, incongruent collocations still give rise to processing costs (Wolter & Gyllstad, 2011, 2013). Studies in which incongruent collocations have been specifically targeted in instruction (Boers, Demecheleer, Coxhead & Webb, 2014, Peters, 2012, 2016) have produced limited gains. Some studies suggest that collocations can be learned incidentally through repeated exposure (Pellicer-Sánchez, 2017; Webb, Newton & Chang, 2013), but enough incidental exposure may not always be possible in contexts of classroom-based learning with limited L2 input. A combination of input flood and text enhancement (e.g., underlining the target collocations) has been shown to lead to greater success in form recognition, but less so in the ability to recall and produce collocations (Szudarski & Carter, 2016). Prior studies have also suggested that input conditions that cause interference should be avoided (Boers et al., 2014, p. 65). While this seems like a

CAMBRIDGE UNIVERSITY PRESS

sensible approach, evidence also indicates that learners experience cross-linguistic interference even if it is not overtly present in the input (Yamashita & Jiang, 2010). Having a high level of cognitive resources or being immersed in the L2 appear to contribute to positive outcomes, but not uniquely (Segalowitz & Freed, 2004; Sunderman & Kroll, 2009). In the study reported here, we take direction from a recent proposal that creating "desirable difficulties" during early stages of learning can result in greater gains (Kornell, Hays & Bjork, 2009; Bjork & Kroll, 2015) to test a new hypothesis about the processing of L2 collocations – one that stems from the observation that classroom-based studies have likely provided appropriate input, but have generally failed to engage the critical cognitive mechanisms that support the learning and retrieval of incongruent collocations. The hypothesis is that to process incongruent L2 collocations efficiently, L2 speakers must learn to inhibit the equivalent L1 collocations, which carry the same meaning, and have the same syntax but only partially overlapping lexical make-up (see Bogulski, Bice & Kroll, 2018 for a similar recent hypothesis about single word learning). These features – syntactic and lexical overlap – have been hypothesized in related L2 research to increase co-activation of L1 collocational competitors. By training lexical selection in conditions that require L1 regulation, learners may be able to develop control mechanisms that are inherent to bilinguals' daily experience of selecting between candidates in two languages. In the approach employed here, we capitalize on previous findings from the memory and cognitive control literature to investigate the representation of incongruent collocations from a novel perspective. Specifically, rather than assuming facilitation due to cross-language congruency (Wolter & Gyllstad, 2011), we examine the prevalence of L1 interference and its effect on lexical selection. We present experimental conditions informed by the findings on the interference from the native language and controlled selection (both discussed below) to train L2 learners' ability to select appropriate representations in an L2.

In the remainder of the introduction, we first discuss current views on the representation and processing of multi-word units (MWUs) in an L2. Next, we discuss the role of the L1 in learning, focusing primarily on classroom-based studies. We then review the literature on interference in L2 memory representations.

1.1. Language non-selective lexical access of multi-word units

While multi-word units are by definition idiomatic and express a holistic meaning, they are also often analyzable into their individual constituents (Bybee, 2010, pp. 25–28; Langacker, 1987). The literature on idioms has provided evidence that the meanings of individual words are also accessed during the processing of idioms (e.g., Gibbs, 1980; Hamblin & Gibbs, 1999; Sprenger, Levelt & Kempen, 2006; Titone & Connine, 1999). Findings show that greater decomposability (i.e., the extent to which the meaning of a multi-word unit can be decomposed based on its individual words) facilitates integration in context when an interpretation must be selected between competing literal or idiomatic meanings (Libben & Titone, 2008; Titone & Libben, 2014). At the same time, there is ample evidence that when idioms can be interpreted literally, there is interference with the non-literal meaning (Caillies & Declercq, 2011; Colombo, 1993). Therefore, while there is evidence of the fine-grained storage of multi-word units in the mind, findings demonstrate that their mental representation often relies on syntactic parsing and on more than one level of semantic representation. Furthermore, the extent to which multi-word units are decomposable and transparent is complicated by the fact that native and non-native speakers may have different perceptions. Specifically, the cultural and linguistic background of learners is likely to affect their perception of aspects such as the transparency of multi-word units (Boers & Webb, 2015).

The fact that multi-word units are analyzable has important implications for L2 learners. It is not only relevant to ask whether conventional expressions in the L1 have equivalents in the L2, but also whether the lexical make-up of such expressions is congruent across languages. In an L2, accessing collocations that are equivalent (i.e., lexically congruent) with the L1 appears to be unproblematic. But recent research has shown that collocations that are incongruent across the L1 and L2 present difficulties in processing, even when they have been learned (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015). While the representation of L2-specific multi-word units has been acknowledged to be problematic for decades (e.g., Irujo, 1984; Nguyen & Webb, 2017; Peters, 2016), the role of previous L1 experience in forming these representations is still not well understood. Wolter and Gyllstad (2011) conducted one of the first studies (see also Yamashita & Jiang, 2010) to systematically investigate the effect of congruency on the processing of L2 collocations. Their results showed that not only were Verb-Noun (V-NP) collocations (e.g., *pay a visit*) processed faster than unrelated V-NP combinations (e.g., *do a visit*), but also that collocations that were congruent across the speaker's two languages were facilitated.

The results were interpreted as an effect of facilitation in processing congruent relative to incongruent collocations. Facilitation accounts assume either "doubled activation" of lexical-semantic nodes across the L1 and L2, or an age-of-acquisition effect due to earlier acquisition of congruent collocations (Wolter & Gyllstad, 2011; Wolter & Yamashita, 2015). An alternative, non-exclusive hypothesis, explored here, is that congruent collocations do not necessarily facilitate online processing; instead, we propose that the implicit activation of L1 lexical links causes interference when L2 speakers process incongruent collocations.

1.2. The role of the L1 in learning L2 collocations

Ample research demonstrates that the degree of similarity between the L1 and the L2 influences the ability to learn and process both the lexicon of a second language (e.g., Brenders, van Hell & Dijkstra, 2011; Dijkstra, Miwa, Brummelhuis, Sappelli & Baayen, 2010; Dijkstra, van Hell & Brenders, 2015) as well as its structure (e.g., Morett & MacWhinney, 2013; Sasaki, 1991; Tolentino & Tokowicz, 2011). Importantly, the presence or absence of L2 structural features in learners' L1 also impacts the effectiveness of the type of instruction provided. What the extant research shows is that L2 features that are absent from learners' native languages require special attention and are best learned when differences are highlighted (Tolentino & Tokowicz, 2014). A relevant perspective in accounting for differences in learning L2-specific structures is the notion of blocking in associative learning (Kamin, 1968), and the idea that entrenched knowledge in the native language may block the formation of L2 representations (for applications of blocking in L2 learning see, e.g., Ellis, Hafeez, Martin, Chen, Boland & Sagarra, 2012; Ellis & Sagarra, 2010).

In what concerns the acquisition of L2 collocations, there is ample evidence that L2 learners often rely on L1 knowledge of how word meanings are combined to learn collocations in the L2 (Biskup, 1992; Bahns, 1993; Irujo, 1984; Granger, 1998;

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

Nesselhauf, 2003, 2005). The discussion in the previous section shows that, relative to congruent collocations, incongruent collocations are at disadvantage in learning and processing. Data from studies on L2 production provide further insight into the role of the L1. A revealing piece of evidence comes from the finding that literal translations (calques) in learners' L2 output that are not idiomatic are often derived from the learners' L1. In a recent study that examined the use of collocations in the free written production of advanced German learners of English, Nesselhauf (2003) described that L1-based translation could explain numerous non-idiomatic choices for the verbs, nouns and prepositions in collocations and idioms. She argued that the evidence of L1 influence, even at higher levels of proficiency, suggests that it is not enough to teach target L2 multi-word units; and that L1 and L2 forms should be explicitly contrasted (Nesselhauf, 2003, p. 239). A classroom-based study by Laufer and Girsai (2008) tested the effectiveness of the L1-L2 contrastive approach in Hebrew-speaking high-school learners of English. In their study, three groups of learners were initially exposed to the same text, and responded to comprehension questions. One day later, one group completed meaning-focused questions that involved using the target collocations; a second group received non-contrastive form-focused questions, including multiple-choice and fill-in exercises; the third group completed questions that involved contrasting and translating from the L1 into the L2 and vice versa. The results of immediate and one-week delayed post-tests showed that the translation treatment in the third group produced the highest rates of learning. Moreover, the rates of learning achieved by the learners in the contrastive treatment seemed to outperform other studies targeting the learning of collocations (although, as the authors acknowledge, comparisons are not always straightforward; Laufer & Girsai, 2008, p. 710). The results from the study suggest that learning can be enhanced through approaches based on contrasting the L1 and L2. However, the results of Laufer and Girsai (2008) provide no insight into the specific mechanisms of retrieval, beyond the fact that an explicit association was established between existing L1 and novel L2 representations. More importantly, it is unclear whether such explicit awareness is directly responsible for the improved performance in the L2, and may facilitate retrieval that is unmediated by contrastive associations with the L1. The focus of the current paper is on investigating the development of the monitoring and selection mechanisms believed to allow for successful and more efficient retrieval and recall.

1.3. Interference in L2 memory representations and the role of cognitive control

Findings from research on phonological, syntactic and semantic priming have provided convincing evidence that access to linguistic representations at different levels is language non-selective (e.g., Carroll & Conklin, 2017; Hartsuiker, Beerts, Loncke, Desmet & Bernolet, 2016; Thierry & Wu, 2007). A consequence is that both languages become activated in parallel, and so L2 speakers must learn to select among competing alternatives available in both their languages. The need to regulate competition is believed to place increased demands on cognitive control in bilinguals (Bialystok, Craik, Green & Gollan, 2009). Moreover, the evidence suggests that resolving competition across languages comes at a cost (e.g., Hoshino & Thierry, 2011). When selecting among competitors, the controlled retrieval of a target candidate renders the alternative not selected more difficult to access. What research on bilingualism has shown is that L2 speakers develop enhanced ability to monitor between conflicting representations (e.g., Abutalebi, 2008), necessary for controlled (rather than automatic) retrieval. We suggest that learners that succeed in learning L1-L2 incongruent constructions are those that can engage control and regulate competition from the L1 (although this discussion concerns lexical learning, for similar arguments regarding general language ability in children see White, Alexander & Greenfield, 2017; L2 proficiency development, Linck, Osthus, Koeth & Bunting, 2014). Our proposal is also informed by the work on paradigms requiring controlled selection, such as the Stroop effect (e.g., name red in the visually-presented word "blue" printed in red), where participants must inhibit related interfering representations in order to selectively control a less accessible response, to train the ability to select the appropriate representations in L2 learners.

2. The present study

This study aims to investigate the association between retrieval conditions and recall of L1-L2 incongruent collocations in an L2. Based on the literature reviewed above, we hypothesize that recall of incongruent collocations is mostly impaired by interference from the native language. As previous research on memory and recall suggests, successfully encoded representations may fail to be retrieved because related information blocks recall (Anderson, Bjork & Bjork, 1994; Bjork & Bjork, 1992). Specifically, we propose that difficulty in recall is due to more strongly active competitors (L1 collocations) blocking retrieval of weakly represented L2 collocations. Therefore, it is expected that, in order to resolve cross-language competition during retrieval, the L1 analogues of target L2 collocations may need to be suppressed. One likely explanation for the limited learning gains reported in classroom-based studies is that learners are not engaging the control mechanisms required to regulate cross-linguistic competition.

A wealth of research on cognitive control indicates that control mechanisms become engaged in conditions that require selection between competing alternatives (e.g., Botvinick, Braver, Barch, Carter & Cohen, 2001; Eriksen & Eriksen, 1974; Stroop, 1935). In lab-based experiments, contexts in which targets are presented alongside distractors allow for the creation of conditions that require participants to select between competitors. Although neutral or unrelated distractors generate low-level conflict, distractors that represent a plausible choice interfere with selection of the target response. We will capitalize on this effect by presenting each of two groups of learners with different practice conditions: one group of learners will see distractors that make plausible verbs based on L1 associations ("L1-Interference group"), while a second group will be presented with unrelated and implausible distractors ("Unrelated group"). As discussed above, collocations are believed to remain analyzable and therefore learners activate the meanings of their individual words. Conflict is hypothesized to emerge through implicit translation in the "L1-Interference" group, in which the distractor verbs will be English translations of the L1 (Spanish) equivalent. Therefore, we will use a between-subject Group manipulation ("L1-Interference" vs "Unrelated" distractors) to examine learning of three different Types of collocation (congruent, incongruent and semantically-related), through performance on two immediate and three delayed recall tests. Our analysis will also consider the contribution of cognitive control by including individual scores derived from the AX-CPT and Flanker tasks, as well as measures of memory from Nonword repetition (PSTM), and Reading Span (WM).

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

2.1. Research questions

In the current study we will aim to address the following questions:

- How does rejection of L2 distractors that are congruent with L1 lexical choices impact lexical selection?
- How does practice in rejecting L1-related distractors while learning L1-L2 incongruent collocations, relative to rejecting unrelated distractors, help regulate L1 interference?
- Can experience with rejecting L1-related distractors aid in the suppression of interference from the native language during subsequent retrieval of L1-L2 incongruent collocations? Will learners who lack practice in suppressing L1-related distractors experience more L1 interference, and therefore produce more errors?
- Will more strongly inhibited L1-related distractors be more costly to subsequently retrieve relative to unrelated distractors that do not require strong inhibition?

2.2. Predictions

Based on the literature, we put forward the following predictions:

1 – As shown in previous research on semantic interference (e.g., Schriefers, 1992), we predict that L1-related distractor verbs will induce cross-language interference, and result in selection costs, producing slower Reaction Times (RTs).
2 – Lexical selection in the L1-Interference group will require the engagement of enhanced monitoring and controlled selection (e.g., Rodriguez-Fornells, van der Lugt, Rotte, Britti, Heinze & Münte, 2005). We predict that the initial cost proposed in Prediction 1 (slower RTs in L1-Interference) will ultimately result in greater efficiency in selection (faster RTs in L1-Interference). This idea is in line with the seemingly paradoxical logic of desirable difficulties, which predicts that apparent initial costs and worse performance in more difficult learning conditions should transform into improved performance and results as learning unfolds.
3 – Practice of controlled retrieval should translate in better access to L2-specific representations in learners in the "L1-Interference" group, leading to enhanced learning relative to the "Unrelated" group. Learning will be measured in recall tests that will allow us to quantify the number of correctly recalled responses, as well as the number of errors that can be attributed to L1 interference (i.e., literal translations congruent with the L1 but not the L2).
4 – If, as predicted, practice in rejecting L1-equivalent distractor verbs helps suppress competition from the L1, those inhibited distractors should become harder to retrieve (slower RTs).

3. Method

3.1. Participants

A group of 49 learners of English was recruited at a university in Spain[1]. Participants were native speakers of Spanish with low-intermediate proficiency in English (levels A2-B1 of the Common European Framework of Reference for Languages, Council of Europe, 2011). All participants gave informed consent and were paid 10 USD per hour of participation. To confirm their eligibility, participants completed a baseline test; eight participants were excluded due to prior knowledge of the collocations. Additional details are provided in the Section 5.3. One participant who reported a learning disability was also excluded. The remaining participants (N = 40) were randomly assigned to one of two learning conditions: a group in which unrelated distractors were presented during the retrieval practice (henceforth, the "Unrelated" group; 75% female, 25% male), and a group that saw L1-related distractors during recall of incongruent collocations ("L1-Interference" group; 70% female, 30% male). The number of familiar items (Unrelated mean: 1.25; SD: 1.48; L1-Interference mean: 0.7; SD: 1.08) was not significantly different between groups ($t(34.8) = -1.3$, $p = .19$).

3.2. Individual differences measures

Participants were administered various measures to assess language proficiency in the L1 and L2, and to ascertain individual differences in cognitive control. Here we summarily enumerate all the measures, and report the results below (see Table 1 below). Complete descriptions of the tasks are available in Appendix A.

To assess linguistic proficiency and background in the L1 and L2, participants completed an abridged version of the LEAP-Q (Marian, Blumenfeld & Kaushanskaya, 2007). General English proficiency was measured through an abridged version of the Michigan English Language Institute College English Test (MELICET; English Language Institute, 2001). Participants also completed a multiple-choice test to assess knowledge of the individual words employed in the experiment, which they were expected to know.

Two measures of memory were collected: (a) A nonword repetition task adapted to Spanish phonotactics was used to measure Phonological Short Term Memory (PSTM) (Baddeley, Papagno & Vallar, 1988; Martin & Ellis, 2012); (b) Working memory was measured through a Spanish version of the Reading Span Task (Elosúa, Gutiérrez, García Madruga, Luque & Gárate, 1996).

Finally, to ensure that both groups were comparable in terms of their cognitive control abilities, participants completed the AX-Continuous Performance Task (AX-CPT) and the Flanker task (Eriksen & Eriksen, 1974).

Groups were matched across all but two measures. A significantly higher mean was found for the Unrelated group in the PSTM task ($t(37.6) = -2.3$, $p = .03$) as well as an earlier L2 Onset of Acquisition (OoA) ($t(32.1) = 2.43$, $p = .02$), relative to the L1-interference group. Both these differences would predict improved learning ability for the Unrelated group relative to the L1-interference group, disfavouring our predictions.

4. Materials

Three types of materials were created: Materials for the Familiarization phase, for the Practice phase and for Testing. The materials for studying and testing were identical for all participants, while the materials used for practice differed across

[1]In determining our target sample size per condition, we aimed for a sample similar to that of other psycholinguistic studies that examined learning conditions under a controlled experimental environment in a laboratory setting (e.g., Finkbeiner & Nicole, 2003; Cintrón-Valentín & Ellis, 2016; Morgan-Short et al., 2014; Potts & Shanks, 2014), including experimental work focused on learning collocations (e.g., Pellicer-Sánchez, 2017; Sonbul & Schmitt, 2013); our sample size was, therefore, guided by the prior literature rather than by a power analysis.

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

Table 1. Summary of cognitive and proficiency measures for each group

	Unrelated condition			L1-interference condition		
	Valid *N*	*M*	*SD*	Valid *N*	*M*	*SD*
Age (in years)	22	24.05	7.5	22	24.65	8.0
Level of education (1–8)	22	4.4	1.2	22	4.2	1.1
OoA (in years)	22	4.6	2.1	22	6.8	3.4
MELICET (/50)	22	20.3	5.1	21	20.8	6.6
Weekly exposure to L2 (/1)	22	0.16	0.1	22	0.17	0.11
L2 Immersion (in years)	22	0.34	0.2	22	0.89	0.3
Eng. Picture Naming (accuracy %)	22	50	10	22	50	10
Span. Picture Naming (accuracy %)	22	95	3.2	22	96	2.5
PSTM: Nonword repetition (/1)	22	0.56	0.10	22	0.48	0.11
WM: Spanish reading span	22	50.35	16	22	46.6	22.9
AX-CPT (BSI)	22	0.40	0.62	22	0.38	0.60
Flanker effect (ms)	22	48.81	24.45	22	48.30	24.04

Note: Means and standard deviations are shown. Values represent raw scores, unless otherwise indicated.

Table 2. Sample of collocations matched across languages and sublists

		L1-L2 incongruent	L1-L2 congruent	Semantically related
1.	English (L2)	*run [a] business*	*carry [his] name*	*walk [the] street*
	Spanish (L1) equivalent	*llevar [un] negocio*	*llevar [su] nombre*	*caminar [la] calle*
	Literal L1 translation	'carry [a] business'	'carry [his] name'	'walk [the] street'
2.	English (L2)	*launder money*	*whiten [one's] teeth*	*clean [one's] hands*
	Spanish (L1) equivalent	*blanquear dinero*	*blanquear [los] dientes*	*limpiar [las] manos*
	Literal L1 translation	'whiten money'	'whiten [one's] teeth'	'clean [one's] hands'

learning groups. All materials were based on the same list of 45 collocations.

For ease of presentation, we first describe the list of collocations used to create the experimental materials. Subsequently, in Section 5, we describe the materials and procedure for the Familiarization, Practice and Testing parts of the experiment.

4.1. List of collocations

Forty-five collocations were extracted using the web-based version of the Corpus of Contemporary American English (COCA, Davies, 2008) with over 560 million words, and data for equivalent Spanish collocations were extracted from the Corpus del Español (Davies, 2016) – with over 2 billion words. Three types of collocations were included (see Table 2 for a sample of the materials). The first type, which we will refer to as "L1-L2 incongruent", differed in Spanish and English by virtue of the fact that the noun was equivalent across both languages but the verb was not. For example, the Spanish equivalent of "run a business" is "llevar un negocio", which literally translates as 'carry a business'. The second type contained "L1-L2 congruent" collocations derived from the first list. To create these, we took the verbs used in the first list ("run" in "run a business"), identified their idiomatic equivalents in Spanish ("llevar un negocio" – literally "carry a business") and used the literal translation to create a new L1-L2 congruent collocation ("carry"; e.g., "carry [his] name"). Each literally translated verb was combined with a new collocational noun that was selected from among the most frequent collocates found in COCA for that verb.

Finally, a third type of "semantically related collocations" also contained congruent collocations, but this time the verbs were semantically related to the verbs of the "incongruent" collocations (e.g., "walk the street", where "walk" is related to "run"). This third type was included to test the ability of participants to learn collocations in which they would experience interference, but from a source different from the L1. At least one study has suggested learning costs when semantically related collocations are learned together (Webb & Kagimoto, 2010). In the present study the meaning of collocations is not related, but the difficulty stems from conflict between semantically related verbs (e.g., walk – run) during recall of the target verb for a given collocation that has been learned. In particular, this manipulation would allow us to examine whether more efficient interference suppression in learners in the L1-interference group might also allow for more efficient selection, despite potential within-language competition in the L2.

Words were not repeated across lists and were matched in word length, frequency, concreteness, and collocational strength. To determine the collocational strength of the verb-noun pairs,

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

Table 3. Testing Protocol

	Week 1		Week 2		Week 3	Week 6
	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6
Tasks	Cognitive battery and proficiency measures	Familiarization and Practice	Practice	Practice	No practice	No practice
Tests		Immediate, after practice	Delayed, at beginning	Immediate, at end of session	1-week delayed test	1-month delayed test

t-scores were used as the statistical association measure. The full stimuli are available in Appendix B.

5. Experimental procedure

Data collection was done over a span of six weeks. In week 1, participants completed the first two sessions of the experiment. In Session 1, participants were administered all the proficiency and cognitive measures. This first session lasted about 2 hours. That same week, participants returned for Session 2. First, they completed the Familiarization phase (lasting approximately 20 min.); this was followed by the experimental Practice procedure (approx. 30 min.), which differed across conditions; and an immediate recall test.

The next two sessions, consisting of additional practice and tests, were completed the following week. The additional Practice sessions and tests allowed us to examine the effect of the manipulation after completing only one Practice session, and after three sessions.

Participants returned for short delayed tests in weeks 3 and 6. Delayed tests were considered critical to assess retention beyond immediate tests, in line with real-life goals of language learning. They should provide a more reliable measure of the long-term impact of the experimental manipulation and the durability of any significant effects observed in immediate testing. An outline of the data collection protocol is presented in Table 3.

5.1. Familiarization phase

Familiarization materials

Familiarization materials were created for auditory and visual presentation of the forty-five English collocations and their Spanish equivalents. The English collocations were recorded by a native English speaker. Spanish collocations were recorded by the first author, a native speaker of the same variety as the speakers tested.

Familiarization procedure

In the Familiarization phase, participants studied the list of 45 collocations, which were presented simultaneously auditorily and visually on a computer screen using E-Prime 2.0 (Schneider, Eschman & Zuccolotto, 2002). This phase served the purpose of familiarizing learners with the target collocations, which would then be practiced under different experimental conditions for each group in the subsequent Practice procedure. Visual presentation of the collocations was done as follows. First, the Spanish collocation was displayed in the center of the screen, followed by a second screen with the English collocation. Participants were then prompted to repeat the English collocation in two subsequent screens, first by saying it out loud, and then by typing it. After a collocation was typed, feedback was provided by showing the collocation in blue or red. If participants typed the collocation exactly as it was shown, the collocation was displayed in blue font; otherwise, the correct collocation appeared in red font (see Figure 1). Overt vocal repetition was used because it has been suggested that it aids learning (e.g., Ghazi-Saidi & Ansaldo, 2017); typed responses allowed to provide automatized feedback during this phase.

Because seeing each of the 45 collocations only once might not provide sufficient familiarization, a repetition round was administered. This time, in addition to repeating the same procedure, participants were asked to recall the Spanish equivalent of the collocations. This was considered important so that learners would not simply focus on learning the form of L2 collocations, while ignoring or failing to remember their meanings. Given the length of the full list (with a total of 45 items), the list was broken down into five segments, with meanings being recalled in blocks of nine collocations.

Oral responses were recorded. Groups of participants did not differ in the ability to recall the meanings of collocations (Unrelated: 85.8%, SD: 35.0; L1-Interference: 85.3%, SD: 35.4; $t(1797.7) = -0.3$, $p = .79$). This Familiarization phase was followed by the Practice procedure. The same Practice procedure was also completed in Sessions 3 and 4 in Week 2, as shown in Table 3. The additional Practice sessions conducted in Week 2 allowed us to examine the effect of the manipulation after completing only one Practice session, and after three sessions at the end of the second week.

5.2. Practice

Practice materials

The materials for the Practice procedure were based on the forty-five collocations studied (e.g., "run a business"). The Practice consisted of forced-choice trials in which the target verb of a collocation had to be selected. Each trial contained two verbs (the target, e.g., "run"; and a distractor verb, e.g., "touch"), and the associated noun in a collocation. In the two experimental learning groups, trials sequentially presented two verbs followed by a noun (V1-V2-N, e.g., "run – touch – business").

One list of Practice materials was created for each of the two experimental learning groups. The two lists differed only in the distractor verbs presented for trials with incongruent collocations. In each list, the incongruent collocations (e.g., "run a business") were associated with distractor verbs specific to each learning group. One group saw unrelated distractors (henceforth the 'Unrelated' group) and another one saw distractors that would be congruent with the native language ('L1-Interference' group). For the previous example ("run a business"), the 'Unrelated' distractor was "touch", while the 'L1-Interference' distractor was "carry"; to reiterate, "carry" was the literal translation of the verb in the Spanish equivalent, as shown in Table 2 above. To

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

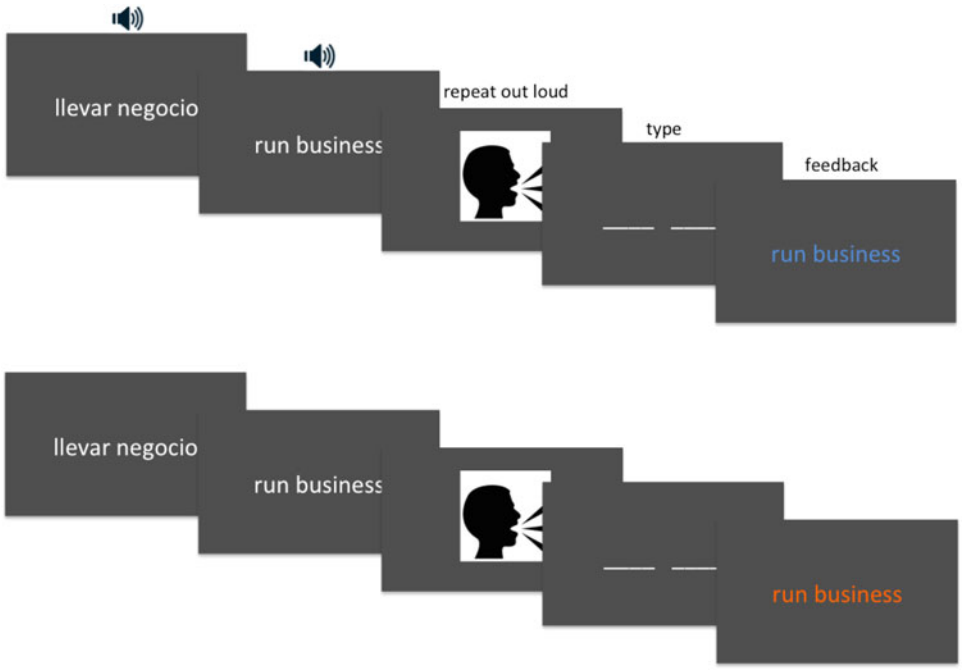

Fig. 1. Sequencing and feedback in the Familiarization procedure

illustrate, participants in the 'Unrelated' group saw e.g., run – touch – business (correct response "run"). For the same collocation, participants in the 'L1-Interference' group had to correctly discard the L1-equivalent distractor verb in e.g., run – carry – business.

The three words in each trial (e.g., run – carry – business) were presented sequentially. The order of target and distractor was counterbalanced, so that the distractor would appear first half of the time, and the target first in the other half. Importantly, all verbs (targets and distractors) are potential candidates, as all are part of the familiarized collocations (i.e., "carry" is part of "carry his name" and "run" appears in "run a business"). Therefore, having seen "run" and "carry" it is not until "business" is displayed that learners could know what verb to select. Each collocation appeared 9 times per list, producing a total of 405 trials per participant in each session.

Practice procedure

Participants were presented with two verbs followed by a noun (V1–V2–N), and were required to respond orally by selecting the appropriate verb accompanying the noun; that is, participants responded by pronouncing the verb selected only. The RTs of the onset of oral responses were automatically recorded by a microphone connected to an SR Box (Psychology Software Tools, Inc.), and accuracy was coded offline. In each trial, a fixation cross was first displayed for 1 second. Then each verb was presented sequentially for 700 ms. The noun was presented for 300 ms, followed by a fixation cross shown for up to 6 seconds or until an oral response was registered, whichever came first. A longer presentation time of the verbs, rather than the noun, was used so that participants would attend to the verbs, and only rely on the noun as a cue for selection between candidates. Feedback was provided by presenting the correct verb for 500 ms after each trial.

5.3. Testing

Testing materials

First, a baseline multiple-choice (MC) test was created in order to assess any potential familiarity with the collocations to be learned, and with the goal of excluding participants with prior knowledge of the target materials. Example 1 below presents a sample item. Each item presented four choices. For the critical incongruent collocations, the four choices contained the target verb (e.g., *run*), the non-target literal Spanish equivalent (*carry*), the associated semantically related verb (*walk*) and a fourth verb (*bring*) in randomized order.

(1) llevar un negocio - _____ a business
(a) run (b) carry (c) walk (d) bring

Participants were asked to provide confidence ratings using a scale from 1 to 5 (1 = no knowledge; 5 = certainty in the response), and those who indicated previous knowledge of at least one collocation (rating = 5) or substantial familiarity with more than three (rating ≥ 3) were excluded (N = 8).

An L1-to-L2 translation test was created to assess immediate and delayed recall at different points throughout the study. For each question, the same Spanish meanings of the English collocations used for the Familiarization phase were presented. Tests were administered in order to assess learning (immediate tests) and retention (delayed tests).

The MC recognition test was used as a baseline rather than a production test, because the inability to recognize the correct verb is a more stringent test of null familiarity than the inability to produce a valid response. That is, our expectation was that learners would not be able to recognize the correct choice when presented to them.

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

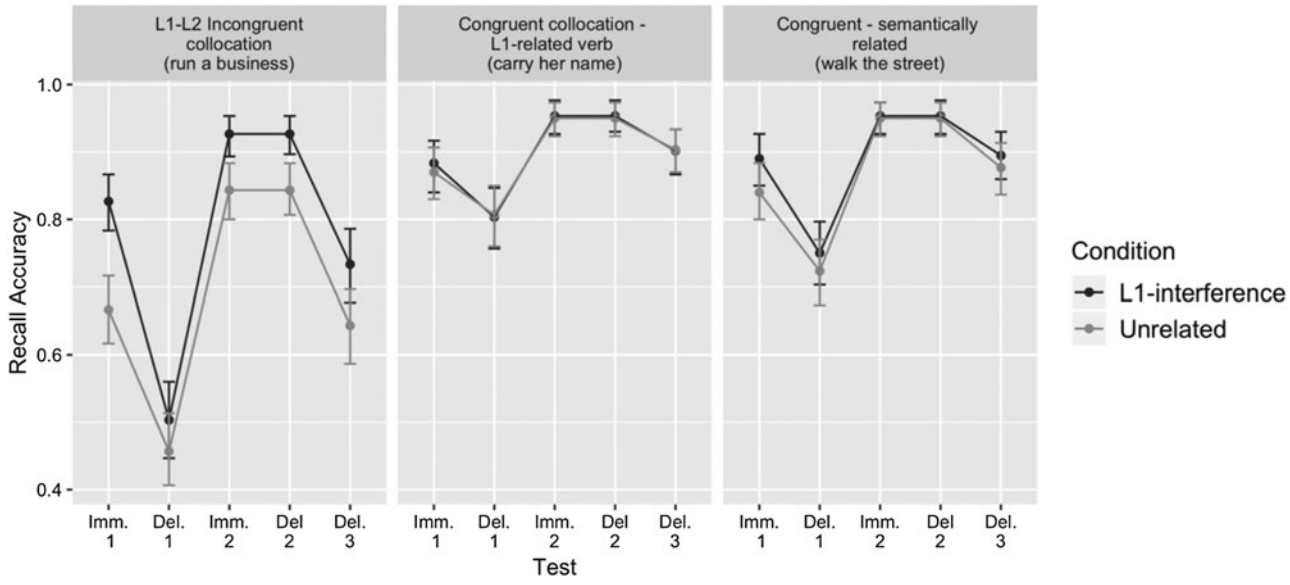

Fig. 2. Results of accuracy rates in immediate and delayed recall tests, for each Type of collocation: Congruent with L1-related verb (left), L1-L2 incongruent (center), and Congruent with semantically related verb (right). Legend: Imm. = Immediate test, Del. = Delayed test. Error bars represent 95% CIs.

Testing procedure

Learners completed immediate tests at the end of the first (Week 1) and last (Week 2) sessions of Practice. The first of three delayed tests was completed in Week 2, while the experiment was still ongoing. Participants returned approximately one week after Practice 1 (mean: 6.2 days; SD: 0.94), and completed the first delayed test before doing additional Practice. The two remaining delayed tests were administered once all Practice was completed. The second and third delayed tests took place approximately one week (mean: 6.8; SD: 1.98) and one month (mean: 31.97; SD: 6.27) after the last Practice session, respectively. One participant in the Interference group did not return for the one-month delayed test; her results are included for all other parts of the study.

For each of the 45 collocations, a screen was presented with a Spanish verb-noun sequence and participants were asked to recall and type the equivalent L2 English collocation. For the immediate tests and the first delayed test, feedback was provided after each response by presenting the verb-noun for 1,500 ms. Since the training procedure involved only recall of verb-noun associations but not of their meaning, feedback on the tests allowed participants to check their form-meaning representations once they had responded. No feedback was provided in the one-week and one-month delayed tests. Responses were coded offline for accuracy.

6. Results of recall tests

In this section, we first present the results of the recall accuracy for the immediate and delayed tests that learners completed. This analysis allowed us to address RQ 3: that is, we examine whether experience rejecting L1-related distractors (L1-Interference group) will result in fewer errors when retrieving L1-L2 incongruent collocations. To further address the question of whether learners in the 'Unrelated' group would experience more interference from the L1, we also conducted an analysis of the types of errors. In section 7, we report RT data for responses in each of the three training sessions.

Typed responses to recall tests were coded for accuracy based on a 2-letter rule. That is, partially correct spellings were accepted as long as no more than 2 characters were misspelled and the correct target word could be identified.

6.1. Accuracy in recall

The results of the recall tests were analyzed using mixed-effects logistic regression. This type of analysis is ideal for binary dependent variables (i.e., accurate or inaccurate response), as it allows one to analyze the unaggregated data rather than means (Jaeger, 2008). All analyses reported in this and other sections were carried out with the lme4 package (Bates, Maechler, Bolder & Walker, 2015) in R version 3.3.2 (R Core Team, 2016).

The analysis examined learning across groups as measured by the tests administered throughout the study (see Figure 2). Following attempts to build maximally specified models (Barr, Levy, Scheepers & Tily, 2013), which led to convergence issues, the random effects structure was simplified (Bates, Kliegl, Vasishth & Baayen, 2015). Final models included random intercepts for subjects and items (Baayen, Davidson & Bates, 2008). In order to control for the baseline differences between groups reported above, models also included by-item slopes for individual PSTM. Fixed effect factors included Group (Unrelated or L1-Interference distractors), Type of collocation (congruent, incongruent or semantically related), Test (each test, 5 levels) and their interaction. The contribution of cognitive control was investigated by including fixed effects for the Behavioral Shift Index (from the AX-CPT), and Flanker Effect. Individual measures of memory were also considered: Nonword repetition (PSTM), and Reading Span (WM). All continuous variables were centered (Baayen et al., 2008).

For each sublist of collocations, models were built starting with a simple mixed-effects structure with Group, Type of collocation, Test, and their interaction, as predictors. Due to convergence issues in models containing the three-way interaction of Group x Type x Test, two-way interactions for Group x Type, Group x Test, and Type x Test were included. In a step-by-step forward model selection procedure, predictors and their interaction with Group and Type were introduced one by one, and were kept if the model fit

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

was significantly improved (likelihood ratio test, $p < .05$). The reference levels were set to Unrelated for Group, and Congruent for Type of collocations. The results are reported below; the model output is available in Appendix C. Parameter-specific p–values were estimated using the normal approximation. Figure 2 shows the progression of performance across groups at each test and for each collocation type.

The analysis revealed a main effect of collocation Type, with significantly lower recall for incongruent collocations in all tests (β: -2.03, SE: 0.58, $p < .001$). There was a marginally significant main effect of Group, with higher accuracy in the 'L1-Interference' group (β: 0.8, SE: 0.45, $p = .07$). Critically, the interaction of L1-Interference Group and incongruent Type was highly significant (β: 0.81, SE: 0.2, $p < .0001$), showing that learners in the 'L1-Interference' group had higher recall rates for incongruent collocations.

The results confirm the pattern shown across testing in Figure 2. Relative to the first immediate test, overall accuracy was significantly lower in the first delayed test (β: -0.62, SE: 0.23, $p < .01$), but higher for tests completed after additional practice (all at least $p < .05$).

However, the interaction of Test x Type showed lower accuracy for incongruent collocations in the Delayed test 1, after only one practice session (β: -0.92, SE: 0.26, $p < .001$); and in the one-month delayed test (β: -0.9, SE:28, $p < .01$). Similarly, the interaction of Test x Group pointed at a reduced effect of Group for Delayed tests 1 (one week after Practice 1; β: -0.58, SE: 0.21, $p < .01$) and 3 (one month after Practice 3; β: -0.49, SE: 0.23, $p < .05$). To further investigate these interactions, we conducted dedicated follow-up analyses with the same methodology described above. The results showed no effect of Group in Delayed test 1, but a marginally significant interaction of Type in the one-month delayed test after completing additional practice (β: 0.8, SE: 0.44, $p = .07$). These results, together with the significant interaction in Delayed 2, revealed that differences in retention were persistent only after practicing recall in the three Practice sessions.

Types of errors

To fully address RQ 3 and explore potential evidence of L1 interference during recall, we conducted an analysis of the types of errors conducted by learners in each group. Specifically, we aimed to explore if learners in the 'Unrelated' group used literal L1 translations more than learners in the 'L1-Intereference' group (who practiced suppression of L1-related translations). In the analysis of types of errors, incorrect test responses that contained an equivalent of the L1 verb (e.g., "carry a business" was produced instead of "run a business") were considered "calques"; the proportion of these was compared against any other type of non-target response (including lack of response). In the Unrelated group, calques accounted for 36% of the errors, but in the L1-Interference group they were only 22% of all errors. A t-test revealed this was a statistically significant difference ($t(292.92) = 2.73$, $p < .01$). The error-type analysis further supports the hypothesis that inducing interference during practice afforded protection against interference in retrieval.

7. Reaction time analysis

This section presents the results of the RT data of responses to trials in the training procedure described above. We first present the results of a Growth Curve Analysis (GCA) on the RTs of incongruent collocations for each individual session, to address RQ 1 and 2 and predictions 1 and 2 formulated above. Questions 1 and 2 were theoretically motivated by the research on desirable difficulties in learning, which predicts that initial costs will transform into advantages. RQ 1 asks whether L1-related distractors will initially cause a cost in selection, i.e., slower RTs; RQ 2 asks whether experience in resolving conflict in responses leads to a more efficient selection process as learning unfolds, relative to the condition with low-conflict ('Unrelated') distractors. We then report the results of a by-group comparison of RTs for different types of collocations in each training session[2], to test RQ 4 and prediction 4.

All analyses were performed on z-score normalized RTs. Trials with invalid RTs due to microphone failure (2%) were removed. Incorrect responses were removed based on offline coding (Unrelated: 7%; L1-Interference: 6%). Responses with RTs shorter than 400 ms after presentation of the noun, or longer than 5,000 ms were excluded (Unrelated: 3%; Interference: 3%).

7.1. Growth curve analysis of reaction time

Growth curve analysis (GCA; Mirman, 2014) was used to analyze the RTs in selection of verbs for incongruent collocations over the course of each of the three Practice sessions. By-session analyses were necessary to determine the shape of the curve within each individual session.

The overall learning curves were modeled with second-order orthogonal polynomials and fixed effects of Group ('Unrelated' or 'L1-Interference') on all time terms. Since the goal of this analysis was to examine selection of the verb only in the incongruent collocations, Type of collocation was not included. Trial order within the training session served as the time dimension. The Unrelated group was treated as the baseline and parameters were estimated for the L1-Interference group. The model also included random effects of participants on all time terms, as well as by-item random slopes for Group, PSTM and OoA. The same fixed effects and interactions as in the accuracy analysis were considered (i.e., PSTM, Flanker Effect, BSI, WM), were added individually and their effects on model fits were evaluated using model comparisons. Parameter-specific p-values were estimated using the normal approximation.

For ease of visual interpretation, the data were binned, as is common practice in GCA (Mirman, 2014, p. 20). In the present analysis, data were grouped into 15 bins; this number of bins maximized the proportion of trials per bin within a range of 10–20 bins. Figure 3 shows the RT data and model fits for the three sessions of training. The model output is available in Appendix C.

Practice session 1

The effect of Group on the intercept did not improve model fit ($\chi2\ (1) = 0.22$; $p = .64$), nor did the effect of Group on the quadratic term ($\chi2\ (1) = 0.95$; $p = .33$). The effect of Group on the linear term, however, improved the model fit ($\chi^2\ (1) = 6.06$; $p < .05$). The interaction of Group with the linear term indicated that the two training conditions differed in the rate of learning, with faster learning in the L1-Interference group (β: -1.09, SE: 0.42, $p < .01$). The results revealed an effect of PSTM (β: -0.05, SE: 0.02, $p < .01$).

[2]Analyses were done by-session, rather than for all at once with session (i.e., time) as a variable. This was done because the significance of some effects might vary from one session to another; by-session analyses helped us avoid potential three- and four-way interactions and allow for easier interpretation of the results.

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

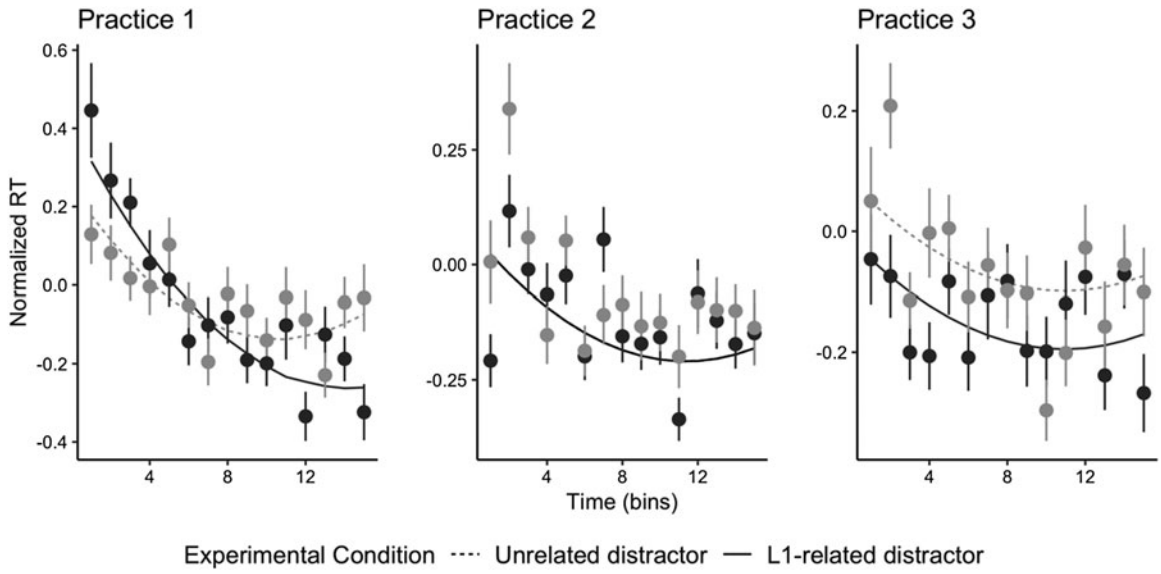

Fig. 3. Growth Curve Analysis of RTs for verb selection across the three practice sessions. For the second practice session, one single solid line represents no significant differences across conditions. All RTs used in the analysis and figures are z-scored. Error bars represent the standard error.

Practice session 2

In the second practice session, Group had no effect on the intercept (χ^2 (1) = 1.45; p = .23), nor on the linear (χ^2 (1) = 0.26; p = .61) or quadratic terms (χ^2 (1) = 1.94; p = .16), suggesting no significant differences across groups.

Practice session 3

No effect of Group was found on the linear (χ^2 (1) = 1.27; p = .26) or quadratic terms (χ^2 (1) = 1.32; p = .25). The model was improved by the effect of Group on the intercept (χ^2 (1) = 4.11; p < .05), showing overall shorter RTs in the L1-Interference group in the session (β: −0.1, SE: 0.04, p < .05).

7.2. Reaction times for verb selection across sublists of collocations

In a final analysis of the RT data, we investigated the cost of selecting the verb for the different lists of collocations across the two learning groups. This allowed us to address RQ 4 and test the associated prediction: that the need for greater inhibition should be associated with a greater cost in retrieval. A mixed-effects linear regression analysis was performed using the same software and procedures described above. In contrast with the GCA reported above, the linear regression analyzed differences within each whole training session. This means that it cannot reveal changes that occur over the course of one session; rather, here we analyze differences in the averages between Groups (Unrelated, L1-Interference) and Types of collocations (Incongruent, Congruent, Semantically related) for the entire session.

Subjects and items were included as random intercepts. The same predictors as in the generalized mixed-effect regression were considered. Final models included by-item random slopes for OoA. Figure 4 shows the RTs of responses in each group for all Practice sessions. Further details are provided in Appendix C.

In comparison with the GCA, models predicting RTs for Sessions 1 and 2 were not improved by adding Group (Session 1: χ^2 (1) = 0.03; p = .86; Session 2: χ^2 (2) = 0.83; p = .66), nor Type of collocation (Session 1: χ^2 (2) = 0.18; p = .91; Session 2: χ^2 (1) = 0.01; p = .92) nor their interaction (Session 1: χ^2 (5) = 1.72; p = .89; Session 2: χ^2 (5) = 3.67; p = .6). However, a significant interaction of Group with Type of collocation was revealed for Session 3, showing that while RTs for verb selection did not differ for congruent collocations across groups, responses for incongruent collocations were faster in the L1-Interference group (β: -0.1, SE: 0.04, p < .05). This result supported our prediction; we discuss the implications of this result in the following section.

There was an interaction between the Flanker effect and collocation type only in Session 1, but not in subsequent sessions. A larger Flanker effect was associated with faster RTs in the incongruent (β: −0.05, SE: 0.02, p < .01) and semantically related (β: −0.06, SE: 0.02, p < .01) lists. Finally, a significant interaction of PSTM with type of collocation emerged for all three sessions, revealing that greater PSTM facilitated selection only in the incongruent and semantically related lists.

We conducted a post-hoc analysis to investigate whether, at the individual level, higher learning gains were associated with L1 inhibition. As an index of interference, we calculated the difference of verb selection RTs in congruent trials minus incongruent trials for the last Practice session (with more positive values indicating inhibition), and correlated it with the scores of the one-month delayed test (Figure 5). A significant correlation for the Unrelated group (r_s = .59, p < .01), confirmed that learners that showed greater inhibition in the L1 were those who tended to have higher recall rates one month later. The analysis of the L1-Interference group, in which L1 inhibition and retention were highest and more generalized, revealed no significant correlation (r_s = .15, p = .53).

8. Discussion

Languages differ in how they encode frequently expressed concepts into conventionalized collocations. Therefore, acquiring a new language also involves learning to combine individual words in ways that are recognizable to other speakers, not only

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

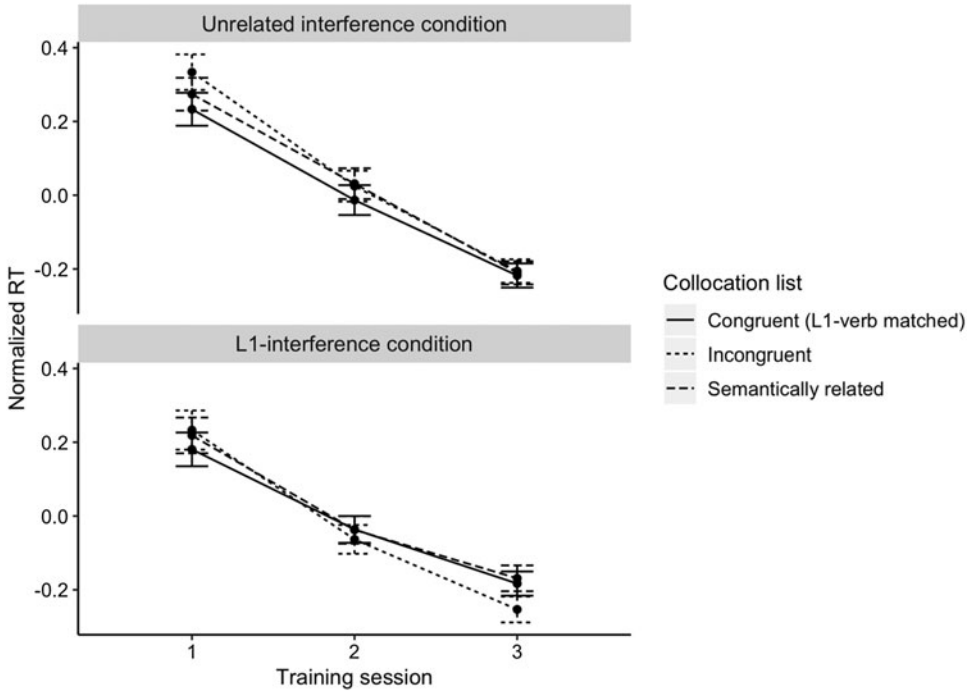

Fig. 4. Whole session RT averages for each of the collocation sublists. The top figure shows RTs in the group of learners with unrelated distractors, and the bottom figure shows the group in the L1-interference condition. All RTs used in the analysis and figures are z-scored. Error bars represent 95% CIs.

in terms of their syntactic structure, but also in their meaning. For adult L2 learners, entrenched knowledge of words and conventional collocations in their native language (e.g., *run a business*) interferes when L2 sequences that are incongruent with those representations must be learned (e.g., *llevar un negocio,* literally 'carry a business' but equivalent to "run a business"; Nesselhauf, 2003; Boers et al., 2014; Peters, 2016). Learners differ in their ability to establish, and then retrieve, these L1-L2 incongruent collocations. We compared learning in two groups to examine if the ability to inhibit competing L1 representations during learning can be trained, and whether this kind of regulation may enhance learning and recall of L2 incongruent collocations. The degree of L1 interference was manipulated in a forced-choice retrieval practice procedure by presenting distractors that were either related or unrelated to the native language equivalents of the sequences being learned.

In a recent study on individual word learning, Bogulski and colleagues (2018) proposed the L1 Regulation Hypothesis, suggesting that learning through L1 translations would engage mechanisms to regulate the native language. The study, which included different groups of bilinguals, demonstrated that those learning through their L1 adopted the strategy of taking longer study time and showed a learning advantage[3]. However, the study did not provide evidence of inhibition of the L1, critical to the regulation hypothesis. The results of the current study showed for the first time that learning conditions that involve the L1, even if implicitly, result in native language inhibition associated with enhanced learning.

8.1. Enhanced learning of incongruent collocations

In line with previous studies (e.g., Peters, 2016), the results of all immediate and delayed recall tests showed that accuracy (i.e., the proportion of correctly recalled items) was significantly lower for incongruent than for congruent collocations in both groups. While this was true for both learning groups, the comparison of recall rates in each group (L1-Interference and Unrelated) revealed that our experimental manipulation was successful in enhancing learning of incongruent collocations in the experimental group (L1-Interference) relative to the baseline (Unrelated). As predicted, inducing interference through L1-related distractors produced significantly greater accuracy for incongruent collocations, even after one single practice session, suggesting that this type of conflict was a desirable difficulty in learning. Further, the testing scheme employed allowed one to assess the amount of practice needed. The results of the first delayed test showed that, after completion of one practice session, between-group differences did not persist when participants were tested one week later. In the second delayed test, one week after two additional practice sessions, recall rates (i.e., the proportion of correctly produced collocations in the L2) remained unchanged. In other words, completing only one practice session produced transient learning, but three practice sessions resulted in stable recall rates. Additionally, the one-month delayed test examined the stability of learning for L2 collocations beyond the two-week period commonly measured in other studies on learning of collocations (e.g., Sonbul & Schmitt, 2013; Boers et al., 2014). When learners were tested again after one month without practice, a marginally significant advantage was still found in accuracy for the

[3]These findings suggest certain advantages in using the native language to develop L1-to-L2 mappings and L1 regulation, and therefore encourages further research into how the L1 may bootstrap L2 learning. The specific pedagogical implications are beyond the scope of this paper. However, recent influential lines of research claim a role for the native language in second language learning in different contexts (e.g., Cummins, 2000) in ways that are distant from the grammar translation methods. The potential benefits of using the L1 in particular contexts are in line with contemporary multilingual pedagogical approaches such as Translanguaging (e.g., Garcia & Wei, 2014; Garcia, Johnson, Seltzer, & Valdés, 2017).

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

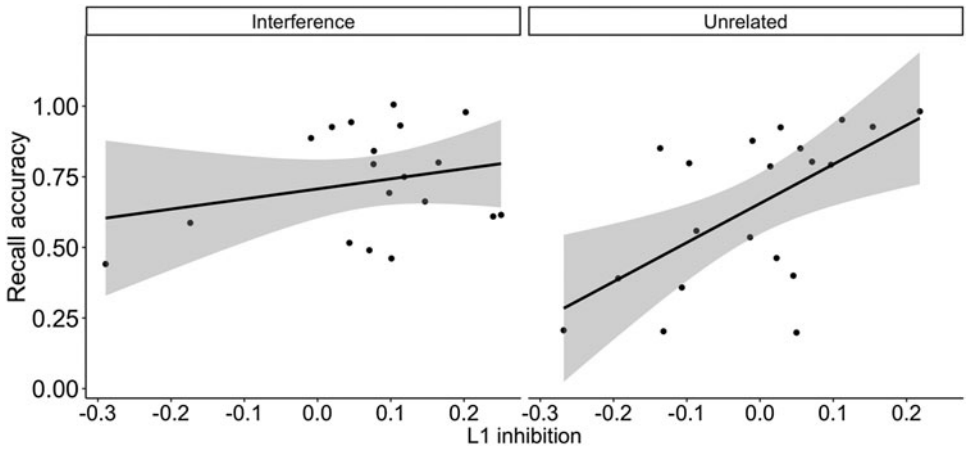

Fig. 5. Correlation of recall rates in the one-month delayed test and index of L1 inhibition. Recall accuracy rates are for the one-month delayed test. The index of L1 inhibition represents the cost of selecting L1-congruent verbs (RTs of L1-equivalent verbs minus RTs of L1-incongruent verbs) in the last practice session. Positive values indicate inhibition of L1-equivalent verbs.

L1-interference group. The test results thus showed that practice in the "desirably difficult" condition, in which L1-related interference was induced, led to increased accuracy in recall of the critical incongruent collocations.

8.2. Native language regulation as part of L2 learning

While the results of the recall tests support our predictions, a main goal of this study was to examine the mechanisms conducive to enhanced recall of collocations. We proposed that induced L1-interference during recall is a desirable difficulty in learning (Bjork & Kroll, 2015; Bogulski et al., 2018), and therefore that initial costs should later lead to greater efficiency in suppressing irrelevant distractors during recall. Specifically, we predicted that if learners who were trained in the L1-Interference condition became more efficient at target selection in incongruent collocations, this should be observable through increasingly faster RTs, and a more rapid rate of learning incongruent collocations. The GCA results were fully consistent with this prediction. The interaction of Group and Type of collocation for RTs in Practice session 1 showed that learners in the L1-Interference group became faster AT A FASTER RATE in selecting the verbs of L1-L2 incongruent collocations than learners in the Unrelated group. The comparison of both groups in Practice session 1 (in Figure 3) shows that initially slower RTs in the L1-Interference group – indicating a greater cost in selection – translated into faster RTs by the end of the training session, relative to the Unrelated group. By Practice session 3, the GCA revealed that RTs in the L1-Interference group were faster throughout the entire session. That is, *harder* retrieval conditions in the L1-Interference group led to *faster* performance than in the Unrelated group during verb selection for incongruent collocations. This supports our claim that conflict-inducing learning conditions engaged control mechanisms that contributed to more efficient selection of incongruent collocations.

An alternative explanation worth considering is that between-group differences were not due to the engagement of mechanisms such as conflict-monitoring and inhibition. Rather, it is possible that presenting L1-related distractors during practice allowed some learners to establish more direct connections between L1 and L2 equivalents. Seeing the verbs of incongruent collocations along with the functionally equivalent verbs of the L1 might have strengthened their association. Indeed, some models of L2 access, such as the Revised Hierarchical Model (Kroll & Stewart, 1994), posit an imbalance in the directionality of lexical connections, with L2-to-L1 connections being weaker than L1-to-L2. Therefore, creating stronger links between the entrenched L1 lexicon and weaker L2 lexicon might aid learning. While this possibility cannot be excluded, a crucial consideration to note is that the implications of strongly associated L2 and L1-equivalent verbs would run counter to the RT data in our results: greater co-activation of equivalent L1 and L2 forms would not result in *less* competition but, quite on the contrary, in GREATER competition from the L1 counterparts. It might well be the case that an association is formed between the target and non-target (i.e., distractor) candidates. Stronger co-activation would result in SLOWER RTs due to conflict in selection; unless the selection system exerts inhibition on the distractors, alleviating competition from highly-active competitors. Therefore, while exposure to verb pairs might facilitate associative links (note that this would apply not just to incongruent collocations, but to all trials), it is unclear how such potential associations might benefit, rather than hamper, conflict resolution among competitors. The only explanation compatible with the finding of faster rejection of distractors in the L1-Interference group, is that selection among competing representations occurs in parallel with increasingly efficient suppression of the non-target candidate.

8.3. Evidence of L1 inhibition during learning

Crucially, the specific prediction that the competitor verbs of incongruent collocations would become inhibited during learning was supported. While both groups of participants performed at ceiling in recalling the verbs of congruent collocations (as expected, given that these congruent collocations are presumed to rely on the L1), emergent costs in retrieval were apparent in the RTs. The results revealed that selection of the verbs of CONGRUENT collocations, which were easiest to recall before training (e.g., *carry* in *carry her name*), became slower after learning the incongruent collocations that require NOT selecting those same verbs (e.g., learning *run a business* requires not selecting the L1-compatible choice *carry a business*). In other words, the verbs that were non-targets for incongruent collocations (e.g.,

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

carry) still had to be selected on some trials as valid targets for congruent collocations; but inhibiting those verbs in incongruent trials (i.e., when they were presented as distractors) resulted in slower RTs for congruent trials where they did need to be selected (i.e., they were the target). This pattern provides strong evidence for the expected inhibition in a specific subset of verbs.

Another piece of evidence supporting the association between L1 regulation and learning comes from within-group variability in the Unrelated condition. While learners in that group had, on average, lower recall rates, some participants did attain high learning outcomes. A significant correlation for the Unrelated group confirmed that learners that showed greater inhibition in the L1 in Practice session 3 were those who tended to have higher recall rates one month later.

These findings point at an association between recall conditions and long-term retention, and suggest that linguistic units that are susceptible to interference during retrieval (potentially resulting in difficult or failed retrieval) benefit from practice conditions that aid development of the necessary interference suppression skills. This idea is the core of theories of learning that posit advantages stemming from conditions that pose desirable difficulties. While the results of recall after one month suggest potential benefits in long-term retention, it is likely that, without any further exposure, gradual memory decay will have a greater long-term impact that outweighs the advantages conferred by more efficient language regulation[4]. Further research will be needed to explore the interaction between language regulation and gradual memory decay in long-term retention (e.g., following months or even years).

8.4. Regulation of lexical competition within the L2

An additional, secondary finding was that not only were the L1-equivalent verbs inhibited, but also the verbs of the congruent collocations in the semantically related sublist (e.g., *walk* is related to *run*), as shown in Figure 4. This should not be surprising, given that within-language inhibition is even more straightforwardly accounted for, fully consistent with the notion that hard-to-retrieve words require inhibition of more highly active competitors. That is, selection of the L1-incongruent verbs resulted not only in the expected inhibition of L1-equivalent verbs (*carry* – *run*), but also in inhibition of within-language semantically related competitors (*walk* – *run*). Given the difficulty in selection of verbs in L1-incongruent collocations, more highly active WITHIN-LANGUAGE semantically-related competitors must also be suppressed. This pattern of results provides insight into effects of both cross- and within-language inhibition in bilinguals.

Finally, a comment regarding competition between different representations is in order. Our stimuli contained collocations along a continuum from more compositional to more idiomatic. As an anonymous reviewer pointed out, the idiom processing literature suggests that idiomatic strings, once acquired, result in significantly faster processing than non-idiomatic strings (e.g., Cacciari & Tabossi, 1988; Libben & Titone, 2008; Titone & Connine, 1999; Titone & Libben, 2014). Incongruent collocations tend to be by definition less cross-linguistically congruent and more idiomatic. Therefore, while incongruent collocations tend to be processed more slowly, less compositional incongruent collocations might show less of a disadvantage. On the other hand, non-native speakers do not always show the same advantages in processing idioms as native speakers (Siyanova, Conklin & Schmitt, 2011), although this may change as proficiency develops (Yeganehjoo & Thai, 2012). However, our materials are not able to directly test this question.

8.5. Types of errors

To further explore potential evidence of inhibition from a different angle, we examined the types of errors produced by learners in the recall tests. A straightforward assumption is that interference from the native language will result in more errors due to use of L1-equivalent verbs (known as "calques"). In this sense, two opposite predictions can be made. The main hypothesis proposed here is that presenting L1-like verb distractors allowed learners to regulate competition from the L1. If this explanation is behind the higher learning rates in the L1-Interference group, we should see, on the one hand, low rates of errors due to L1 intrusions in that condition. On the other hand, learners in the Unrelated group – even though they were not presented with L1-related choices – would be predicted to produce more calques. Such a pattern would be further evidence that errors are due to the lack of experience in suppressing interference.

The alternative approach, discussed above, is that presenting the L1-equivalent verbs of incongruent collocations simply reinforced L1-L2 associations, and better recall was not due to better monitoring of interfering representations, but to facilitating cross-language associations. This would lead to very different predictions: because learners in the Unrelated group never saw the L1-congruent verb paired with the L2-target verb, L1-based calques should have been rare. On the other hand, learners in the L1-Interference group, who repeatedly saw the L1 equivalent, might produce a high number of L1-congruent intrusions. The error-type analysis reported provides further support to the hypothesis that inducing interference during practice afforded protection against interference in retrieval.

8.6. Individual differences in memory

The role of two different memory constructs – namely, PSTM and WM – was examined by including these measures as fixed effects in the analyses performed. The results showed the paramount role of phonological memory in learning of collocations, in the line of similar findings in previous studies of single word learning (Kaushanskaya, 2012; Martin & Ellis, 2012). The main effect of PSTM in every recall test points to the unique contribution of phonological memory to the encoding and retrieval of collocations as unitary chunks. WM was only significant in the second immediate test, administered after all three sessions of retrieval practice were completed, but not in the other tests.

It is worth noting that the second immediate test was completed at the point in which the benefits of practice were maximal. The significant contribution of WM in this particular test suggests an important role in allowing retrieval of the correct verb-noun combinations at a time in which all competitors were maximally active. We suggest that the ability to resolve competition among simultaneously active candidates may be the mechanism behind the significant contribution of WM.

[4]We thank one anonymous reviewer who asked about the potential consequences of L1 inhibition for attrition of the native language. While this point will need to be addressed in further research, we would like to argue that language attrition is not necessarily only by L1 inhibition, and requires also decreased usage of the L1.

https://doi.org/10.1017/S1366728919000622 Published online by Cambridge University Press

8.7. The role of cognitive control in learning incongruent collocations

The contribution of cognitive control was investigated by considering the measures from Flanker task and AX-CPT in the analyses. Because retrieval of target L2 collocations was dependent on inhibition of L1 competitors, we predicted a significant role of cognitive control. The results showed that the Flanker effect scores (RTs in incongruent trials – RT in congruent trials) were predictive of RT group averages in the first practice session. Somewhat counterintuitively, a greater Flanker effect (indicating less efficient inhibition) was associated with faster RTs in selection of the target verbs in incongruent collocations during retrieval practice 1 (β: -0.05, SE: 0.02, $p < .01$); no effects were found for practice 2 and 3.

As discussed, better learning was associated with a greater initial cost in selection, shown by longer RTs in Practice 1 (see Figure 4). That is, greater learning was associated with SLOWER RTs, which reflected that conflict-monitoring mechanisms were in fact being engaged appropriately. In LESS SUCCESSFUL learners, RTs were FASTER, as shown by the results for the Unrelated group. Thus, the finding that those learners with greater ability to detect conflict experienced greater delays in response selection is consistent with the notion that enhanced learning was associated with a greater initial cost, and that experiencing greater interference was associated with greater gains when L2 representations are in conflict with the L1.

8.8. Generalizability of efficient learning

In addition to investigating the effect of distractors on recall of the incongruent collocations, we asked whether greater ability to select among competitors in some learners might translate into an advantage when resolving other types of interference. To this end, a subgroup of the congruent collocations was included, whose verbs were semantically related to the verbs of the incongruent collocations (e.g., run – walk). A greater cost of selection in this sublist of congruent collocations was observed in a number of ways. First, the analysis of whole session RTs showed that individual PSTM predicted speed in verb selection not just for the incongruent collocations but also for collocations with semantically related verbs. Moreover, in Practice 1, individual Flanker effect scores also predicted RTs in verb selection for the incongruent and semantically related sublists, but not for the other congruent collocations. Despite the cost from semantic interference reflected in RTs, the lack of interactions suggested there were no differences across groups. A limitation to be noted is that recall of congruent collocations (including the semantically related) produced ceiling effects. Future studies should further investigate the potential spillover from enhanced efficiency in learning by examining more challenging learning conditions that prevent ceiling effects.

Finally, we can speculate that a desirable difficulties-based approach might confer other advantages in learning. However, desirable difficulties in learning e.g., tense morphology, may focus on mechanisms other than interference. For instance, learning of morphological cues specific to the L2 may require avoiding reliance on known cues that also exist in the L1 (e.g., see Cintrón-Valentín & Ellis, 2016; Ellis et al., 2012).

9. Conclusion

The current study provides new evidence of the weight that regulating competition from the native language carries in learning collocations in another language. Moreover, it tested a new practice paradigm focused on suppression of L1-interference as a new type of desirable difficulty in language learning. The results provided evidence that retrieval practice in interference-induced conditions enhanced learning gains of incongruent collocations relative to the baseline (unrelated-distractor) group. Further, the association between L1-inhibition and L2 learning gives critical support to the L1 Regulation hypothesis (Bogulski et al., 2018). Future research should further investigate how engaging cognitive processes through particular learning conditions may produce not just better learning outcomes but also better learners. The findings reported also have practical implications for learning and teaching pedagogy, and for the design of learner-oriented tests and materials.

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728919000622

Acknowledgements. The authors would like to thank the ILCE at the Universidad de Navarra and Dr. Eduardo Negueruela for providing the facilities to conduct this research; Dr. Mª Rosa Elosúa for sharing the materials for the Spanish reading span task; Christian Navarro-Torres and Dr. Laura Rodrigo for their help with GCA modeling; Paul Bentham for recording the English stimuli; Daniel Ng and Anna Northcott for help with coding; and the associate editor and three anonymous reviewers for their input. Any errors remain our own. The writing of this article was supported in part by National Science Foundation grants BCS-1535124 and OISE-1545900 to Paola E. Dussias.

References

Abutalebi J (2008) Neural aspects of second language representation and language control. *Acta Psychologica* **128**, 466–478. doi:10.1016/j.actpsy.2008.03.014

Anderson MC, Bjork RA and Bjork EL (1994) Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **20**, 1063–1087. doi:10.1037/0278-7393.20.5.1063

Arnon I and Snider N (2010) More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* **62**, 67–82. doi:10.1016/j.jml.2009.09.005

Baayen RH, Davidson DJ and Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**, 390–412. doi:10.1016/j.jml.2007.12.005

Bahns J (1993) Lexical collocations: A contrastive view. *ELT Journal* **47**, 56–63. doi:10.1093/elt/47.1.56

Baddeley A, Papagno C and Vallar G (1988) When long-term learning depends on short-term storage. *Journal of Memory and Language* **27**, 586–595. doi:10.1016/0749-596X(88)90028-9

Bardovi-Harlig K (2009) Conventional expressions as a pragmalinguistic resource: Recognition and production of conventional expressions in L2 pragmatics. Language Learning: *A Journal of Research in Language Studies* **59**, 755–795. doi:10.1111/j.1467-9922.2009.00525.x

Bates D, Maechler M, Bolker B and Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48. doi:10.18637/jss.v067.i01.

Bates D, Kliegl R, Vasishth S and Baayen H (2015) Parsimonious mixed models. arXiv preprint arXiv:1506.04967

Barr DJ, Levy R, Scheepers C and Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* **68**, 255–278. doi:10.1016/j.jml.2012.11.001

Bialystok E, Craik FIM, Green DW and Gollan TH (2009) Bilingual minds. *Psychological Science in the Public Interest* **10**, 89–129. doi:10.1177/1529100610387084

Biskup D (1992) L1 influence on learners' renderings of English collocations: A Polish/German empirical study. In Arnaud P. and Bejoint H. (eds.), *Vocabulary and applied linguistics*. Amsterdam: John Benjamins, pp. 85–93.

Bjork RA and Bjork EL (1992) A new theory of disuse and an old theory of stimulus fluctuation. In Healy A, Kosslyn S., & Shiffrin R. (eds.), *From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes*. Hillsdale, NJ: Erlbaum, **2**, pp. 35–67.

Bjork RA and Kroll JF (2015) Desirable difficulties in vocabulary learning. *The American Journal of Psychology* **128**, 241–252. doi:10.5406/amerjpsyc.128.2.0241

Boers F, Eyckmans J, Kappel J, Stengers H and Demecheleer M (2006) Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research* **10**, 245–261. doi:10.1191/1362168806lr195oa

Boers F, Demecheleer M, Coxhead A and Webb S (2014) Gauging the effects of exercises on verb–noun collocations. *Language Teaching Research* **18**, 54–74. doi:10.1177/1362168813505389

Boers F and Webb S (2015) Gauging the semantic transparency of idioms: do natives and learners see eye to eye? In Heredia RR and Cieślicka AB (eds.), *Bilingual figurative language processing*. New York, NY: Cambridge University Press, pp. 368–392.

Bogulski CA, Bice K and Kroll JF (2018) Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition* 1–16. doi:10.1017/S1366728918000858

Botvinick MM, Braver TS, Barch DM, Carter CS and Cohen JD (2001) Conflict monitoring and cognitive control. *Psychological Review* **108**, 624–652.

Brenders P, van Hell JG and Dijkstra T (2011) Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology* **109**, 383–396. doi:10.1016/j.jecp.2011.03.012

Bybee JL (2010) *Language, usage and cognition*. Cambridge; New York: Cambridge University Press.

Cacciari C and Tabossi P (1988) The comprehension of idioms. *Journal of Memory and Language* **27**, 668.

Caillies S and Declercq C (2011) Kill the song – steal the show: What does distinguish predicative metaphors from decomposable idioms? *Journal of Psycholinguistic Research* **40**, 205–223. doi:10.1007/s10936-010-9165-8

Carrol G and Conklin K (2017) Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'. *Bilingualism: Language and Cognition* **20**, 299–317. doi:10.1017/S1366728915000103

Cintrón-Valentín MC and Ellis NC (2016) Salience in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology* **7**, 1284.

Colombo L (1993) The comprehension of ambiguous idioms in context. In Cacciari C. and Tabossi P. (eds.), *Idioms: Processing, structure, and interpretation*. Hillsdale, NJ: Erlbaum, pp. 163–200.

Council of Europe (2011) Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Council of Europe.

Cummins J (2000) *Language, power, and pedagogy: Bilingual children in the crossfire* (vol. 23). Multilingual Matters.

Davies M (2008) *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. https://corpus.byu.edu/coca/

Davies M (2016) *Corpus del Español: Two billion words, 21 countries*. http://www.corpusdelespanol.org/web-dial/

Dijkstra T, Miwa K, Brummelhuis B, Sappelli M and Baayen H (2010) How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language* **62**, 284–301. doi:10.1016/j.jml.2009.12.003

Dijkstra T, Hell JGv and Brenders P (2015) Sentence context effects in bilingual word recognition: Cognate status, sentence language, and semantic constraint. *Bilingualism: Language and Cognition* **18**, 597–613. doi:10.1017/S1366728914000388

Durrant P and Doherty A (2010) Are high-frequency collocations psychologically real? investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory* **6**, 125–155.

Ellis NC, Hafeez K, Martin KI, Chen L, Boland J and Sagarra N (2012) An eye-tracking study of learned attention in second language acquisition. *Applied Psycholinguistics* **35**, 547–579.

Ellis NC and Sagarra N (2010) The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition* 32, 552–580.

Elosúa MR, Gutiérrez F, García Madruga JA, Luque JL and Gárate M (1996) Adaptación española del reading span test de Daneman y Carpenter. *Psicothema* **8**, 383–395.

English Language Institute. (2001) *MELICET – GCVR User's Manual*. Ann Arbor, MI.

Eriksen BA and Eriksen CW (1974) Effects of noise letters upon the identification of target letters in a nonsearch task. *Perception & Psychophysics* **16**, 143–149.

Finkbeiner M and Nicol J (2003) Semantic category effects in second language word learning. *Applied psycholinguistics* **24**, 369–383

García O and Wei L (2014) Translanguaging. *The Encyclopedia of Applied Linguistics* 1–7.

García O, Johnson SI, Seltzer K and Valdés G (2017) *The translanguaging classroom: Leveraging student bilingualism for learning*. Philadelphia, PA: Caslon.

Ghazi-Saidi L and Ansaldo AI (2017) Second language word learning through repetition and imitation: Functional networks as a function of learning phase and language distance. *Frontiers in Human Neuroscience* **11**, 463–463. doi:10.3389/fnhum.2017.00463

Gibbs Jr. RW (1980) Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition* **8**, 149–156.

Granger S (1998) Prefabricated patterns in advanced EFL writing: Collocations and formulae. In Cowie AP (ed.), *Phraseology: Theory, analysis, and applications*. Oxford, UK: Oxford University Press, pp. 145–160.

Gries ST (2013) 50-something years of work on collocations. *International Journal of Corpus Linguistics* **18**, 137–165.

Hamblin JL and Gibbs Jr. RW (1999) Why you can't kick the bucket as you slowly die: Verbs in idiom comprehension. *Journal of Psycholinguistic Research* **28**, 25–39.

Hartsuiker RJ, Beerts S, Loncke M, Desmet T and Bernolet S (2016) Cross-linguistic structural priming in multilinguals: Further evidence for shared syntax. *Journal of Memory and Language* **90**, 14–30. doi:10.1016/j.jml.2016.03.003

Hoshino N and Thierry G (2011) Language selection in bilingual word production: Electrophysiological evidence for cross-language competition. *Brain Research* **1371**, 100–109. doi:10.1016/j.brainres.2010.11.053

Ibbotson P (2013) The scope of usage-based theory. *Frontiers in Psychology* **4**, 255. doi:10.3389/fpsyg.2013.00255

Irujo S (1984) Don't Put Your Leg in Your Mouth: Transfer in the Acquisition of Idioms in a Second Language. Presented at the Boston University Conference on Language Development, Boston.

Jaeger TF (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* **59**, 434–446. doi:10.1016/j.jml.2007.11.007.

Kamin LJ (1968) "Attention-like" processes in classical conditioning. In Jones M. R. (ed.), *Miami symposium on the prediction of behavior: Aversive stimulation*. Coral Gabels, FL: University of Miami Press, pp. 9–31.

Kaushanskaya M (2012) Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition* **15**, 1–20. doi:10.1017/S1366728911000472

Kornell N, Hays MJ and Bjork RA (2009) Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **35**, 989–998. doi:10.1037/a0015729

Kroll JF and Stewart E (1994) Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* **33**, 149–174. doi:10.1006/jmla.1994.1008

Langacker R (1987) *Foundations of cognitive grammar*. Stanford University Press.

Laufer B and Girsai N (2008) Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics* **29**, 694–716. doi:10.1093/applin/amn018

Libben MR and Titone DA (2008) The multidetermined nature of idiom processing. *Memory & Cognition* **36**, 1103–1121. doi:10.3758/MC.36.6.1103

Linck JA, Osthus P, Koeth JT and Bunting MF (2014) Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review* **21**, 861–883. doi:10.3758/s13423-013-0565-2

Marian V, Blumenfeld HK and Kaushanskaya M (2007) The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* **50**, 940–967. doi:10.1044/1092-4388(2007/067)

Martin KI and Ellis NC (2012) The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition* **34**, 379. doi:10.1017/S0272263112000125

Mirman D (2014) *Growth curve analysis and visualization using R*. Boca Raton: CRC Press.

Morett LM and MacWhinney B (2013) Syntactic transfer in english-speaking spanish learners. *Bilingualism: Language and Cognition* **16**, 132–151. doi:10.1017/S1366728912000107

Morgan-Short K, Faretta-Stutenberg M, Brill-Schuetz KA, Carpenter H and Wong PC (2014) Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition* **17**, 56–72.

Nesselhauf N (2003) The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics* **24**, 223–242. doi:10.1093/applin/24.2.223

Nesselhauf N (2005) *Collocations in a learner corpus*. Amsterdam; Philadelphia: J. Benjamins Publishing Company.

Nguyen TMH and Webb S (2017) Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research* **21**, 298–320. doi:10.1177/1362168816639619

Pellicer-Sánchez A (2017) Learning L2 collocations incidentally from reading. *Language Teaching Research* **21**, 381–40.

Peters E (2012) Learning german formulaic sequences: The effect of two attention-drawing techniques. *The Language Learning Journal* **40**, 65–79. doi:10.1080/09571736.2012.658224

Peters E (2016) The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research* **20**, 113–138. doi:10.1177/1362168814568131

Potts R and Shanks DR (2014) The benefit of generating errors during learning. *Journal of Experimental Psychology: General* **143**, 644.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rodriguez-Fornells A, Lugt Avd, Rotte M, Britti B, Heinze H and Münte TF (2005) Second language interferes with word production in fluent bilinguals: Brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience* **17**, 422–433. doi:10.1162/0898929053279559

Sasaki Y (1991) English and Japanese interlanguage comprehension strategies: An analysis based on the competition model. *Applied Psycholinguistics* **12**, 47–73.

Segalowitz N and Freed BF (2004) Context, contact, and cognition in oral fluency acquisition: Learning spanish in at home and study abroad contexts. *Studies in Second Language Acquisition* **26**, 173–199. doi:10.1017/S0272263104262027

Siyanova-Chanturia A, Conklin K, Caffarra S, Kaan E and van Heuven WJB (2017) Representation and processing of multi-word expressions in the brain. *Brain and Language* **175**, 111–122.

Siyanova-Chanturia A, Conklin K and Schmitt N (2011) Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research* **27**, 251–272. doi:10.1177/0267658310382068

Schneider W, Eschman A and Zuccolotto A (2002) *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools, Incorporated.

Schriefers H (1992) Lexical access in the production of noun phrases. *Cognition* **45**, 33–54. doi:10.1016/0010-0277(92)90022-A

Sonbul S and Schmitt N (2013) Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions: Explicit and implicit lexical knowledge. *Language Learning* **63**, 121–159. doi:10.1111/j.1467-9922.2012.00730.x

Sprenger SA, Levelt WJM and Kempen G (2006) Lexical access during the production of idiomatic phrases. *Journal of Memory & Language* **54**, 161–184.

Stroop JR (1935) Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* **18**, 643–661.

Sunderman G and Kroll JF (2009) When study-abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics* **30**, 79–99. doi:10.1017/S0142716408090048

Szudarski P and Carter R (2016) The role of input flood and input enhancement in EFL learners' acquisition of collocations: L2 input types and acquisition of collocations. *International Journal of Applied Linguistics* **26**, 245–265. doi:10.1111/ijal.12092

Thierry G and Wu YJ (2007) Brain potentials reveal unconscious translation during foreign language comprehension. *Proceeding of National Academy of Sciences* **104**, 12530–12535.

Titone DA and Connine CM (1999) On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics* **31**, 1655–1674.

Titone DA and Libben MR (2014) Time-dependent effects of decomposability, familiarity and literal plausibility on idiom meaning activation: A cross-modal priming investigation. *The Mental Lexicon* **9**, 473–496. doi:10.1075/ml.9.3.05tit

Tolentino LC, & Tokowicz N (2011) Across languages, space, and time: A review of the role of cross-language similarity in L2 (morpho)syntactic processing as revealed by fMRI and ERP methods. *Studies in Second Language Acquisition* **33**, 91–125.

Tolentino LC and Tokowicz N (2014) Cross-Language similarity modulates effectiveness of second language grammar instruction. *Language Learning* **64**, 279–309. doi:10.1111/lang.12048

Webb S and Kagimoto E (2010) Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics* **32**, 259–276.

Webb S, Newton J and Chang A (2013) Incidental learning of collocation. *Language Learning* **63**, 91–120. doi:10.1111/j.1467-9922.2012.00729.x

White LJ, Alexander A and Greenfield DB (2017) The relationship between executive functioning and language: Examining vocabulary, syntax, and language learning in preschoolers attending head start. *Journal of Experimental Child Psychology* **164**, 16–31. doi:10.1016/j.jecp.2017.06.010

Wolter B and Gyllstad H (2011) Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics* **32**, 430–449. doi:10.1093/applin/amr011

Wolter B and Gyllstad H (2013) Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition* **35**, 451. doi:10.1017/S0272263113000107

Wolter B and Yamashita J (2015) Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics* **36**, 1193–1221. doi:10.1017/S0142716414000113

Yamashita J and Jiang N (2010) L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly* **44**, 647–668.

Yeganehjoo M and Thai YN (2012) Lexical access in production of idioms by proficient L2 learners. *The Southeast Asian Journal of English Language Studies* **18**, 87–104.