Cross-linguistic Collocational Networks in the L1 Turkish - L2 English Mental Lexicon

*Dr. Hakan CANGIR, Ankara University**
Dr. Philip Durrant, University of Exeter

ABSTRACT

According to the *Collocational Priming Theory*, every word is primed to co-occur with particular other related words and priming could be regarded as the source of our creative language system. Previous research has shown evidence of collocational priming in both L1 and L2 users of English and has indicated that L2 processing is influenced by L1 collocations. This study attempts to further our understanding of the relationship between first and second language collocations through the paradigm of *cross-linguistic priming*. That is, it will test the extent to which individual words in one language prime recognition of those words' collocates in the other language. Results suggest a complex picture of both cross-linguistic priming and cross-linguistic inhibition, operating differently across different part of speech combinations. They also suggest important methodological influences which future research will need to investigate. Findings are discussed in the light of the current bilingual mental lexicon models and some implications are drawn based on the observed collocational networks in the L1 Turkish-L2 English bilingual mental lexicon.

Key Words: Collocational priming, Cross-linguistic, Frequency, Mental Lexicon

1. INTRODUCTION

It is widely agreed among researchers that collocations, which are defined here as frequently co-occurring word combinations, are pervasive in language (Schmitt, 2010). Hoey (2005) argued that the psycholinguistic mechanism of *priming* could be used to explain how language users come to learn and successfully produce such collocations. In this model, first proposed by Meyer and Schvaneveldt (1971), a language user's recognition of a word is facilitated (or inhibited) by the words they have recently encountered. Specifically, words are recognized more rapidly if they occur soon after words to which they are related in some way. Thus, *doctor* is recognized more quickly if it is encountered after *nurse* than it would be if encountered after a less closely-related word. In this case, the word *nurse* is said to *prime* the target word *doctor.* Priming is interpreted as giving us insight into the structure of the mental lexicon. Specifically, psycholinguists have suggested that priming between words is due to neurological activity *spreading* from the priming word to the primed word (Collins and Loftus, 1975).

Hoey's model applies this notion to collocation, suggesting that speakers choose high-frequency collocates over plausible alternatives because, when a particular word is brought

*Corresponding author: Hakan CANGIR, Ankara University, School of Foreign Languages, Turkey, E-mail: hcangir@ankara.edu.tr
Philip DURRANT, University of Exeter, Graduate School of Education, Exeter, E-mail: P.L.Durrant@exeter.ac.uk

to mind (e.g. *rain*), recall of its collocates (e.g. *heavy*) is facilitated, thus making it more likely that those collocating words will be chosen over contending words with similar meanings (e.g. *strong*).

Research into priming between collocating words has lent some empirical support to Hoey's model, providing evidence for an effect in both first (e.g., Durrant and Doherty, 2010) and second (e.g., Wolter and Gyllstad, 2011; Yamashita, 2018) language users. However, an interesting outstanding question concerns how collocation operates *across* languages. That is, the extent to which collocations in a speaker's first language influences their processing of collocations in a second language, and vice-versa. That such an effect is likely is strongly suggested by the wealth of studies demonstrating that second language knowledge and use of collocations is influenced by the collocations of the first language (e.g., Yamashita and Jiang, 2010; Zinkgraf, 2008; Huang, 2001). Psycholinguistic studies have also demonstrated that second language collocations are processed faster when they are *congruent* with collocations in the first language; that is, where the usual second language collocation is a word-for-word translation of a collocation in the first language (Wolter and Gyllstad, 2011; 2013; Wolter and Yamashita, 2014). In one of the few studies scrutinzing the processing of formulaic units at a cross-linguistic level, Carrol and Conklin (2017) conducted two eye-tracking experiments investigating cross-language lexical priming in formulaic units. They also concluded that congruence in formulaic language facilitated cross-language processing and that a cross-linguistic interaction seemed to exist at the multiword level.

What these studies do not address, however, is the question of *cross-linguistic collocational priming*: that is, the extent to which a prime word (a node) in the first language influences the processing of a target word (a collocate) in the second language, or vice-versa. Cross-linguistic priming effects between semantically-related words (e.g. *pencil* in English priming *goma*, the Spanish word for eraser) and translation equivalents (e.g. *ball* in English priming *balon* in Spanish) are widely-attested in the literature, and play an important role in models of the bilingual lexicon (see Altarriba and Basnight-Brown, 2007 for an overview).

One of the most consistent findings of cross-language priming research is an asymettry in cross-language lexical processing. Some studies conclude that we can observe a priming effect only from the dominant language to the less dominant language. Altarriba (1992), for instance, detected translation priming effects in this direction in an experiment with a short stimulus onset asynchrony (SOA). Keatley et al. (1994), whose participants were Chinese-English and French-Dutch bilinguals, observed the same priming effect (only from L1 to L2) for associative prime-target pairs. A more recent study (Kim and Davis, 2003) also stated that there was a cross-language priming effect in noncognate translation pairs in the L1-L2 direction. However, there are also some conflicting results in the literature. For instance, Basnight-Brown and Altarriba (2007) found masked priming effects in both directions in Spanish-English bilinguals. Duyck et al. (2008) detected the same effect for translation priming in L2-L1 and L2-L3. These contradictory findings may be related to some methodological alternations and to the language backgrounds of the participants and should be considered very carefully while setting up the experimental design (Jiang, 2015).

Taken together, the existing liteature provides a large body of knowledge about cross-linguistic lexical priming and has enabled us to understand many of factors that impact on the processing of lexical items cross-linguistically. It allows us to understand how such priming varies depending on different experimental tasks (e.g. naming, lexical decision and semantic categorization); the status of words as cognates or noncognates; different proficiency levels of participants; different language pairs; and different prime word durations. However, the ways in which collocations are processed cross-linguistically has, to date, gone unresearched. This will be the focus of the current study.

1.1. The Current Study

The current study aims to provide initial evidence on the occurrence of cross-linguistic priming between Turkish and English in first language speakers of Turkish who are fairly proficient in English. Turkish makes an interesting comparison with English due to its strongly agglutinative character, which leads to corpus frequency profiles that are distributionally very different from those found in English (Durrant, 2013; Öksüz, 2019). Despite these differences, previous research has suggested both that first language collocational priming exists within Turkish (Cangır, Büyükkantarcıoğlu and Durrant, 2017; Öksüz et al., 2020) and that first language speakers of Turkish with a good knowledge of English show priming in English despite the strong typological differences between those two language.

Thus, our first research question is:

1. Does cross-linguistic collocational priming occur between Turkish and English for L1 Turkish and L2 English users?

The existing literature on second language collocational priming suggests that effects may differ depending on the distributional properties of collocations (frequency and strength of association) and congruence between first and second language collocations (Wolter and Gyllstad, 2011; 2013; Wolter and Yamashita, 2014; Yamashita, 2018). Moreover, the cross-linguistic priming literature has shown that effects tend to be stronger in the direction of first language prime to second language target than vice-versa (Altarriba and Basnight-Brown, 2007). It is also intuitively likely that any cross-linguistic collocational effects between English and Turkish will be partly dependent on the syntactic structure of the particular combinations studied. The effects of agglutination in Turkish are seen most strongly on the verb, which takes a range of concatenating suffixes to show, amongst other things, tense, aspect, person, voice, modality, causality, evidentiality, etc. (Durrant, 2013). This results in many different realisations of each verb, each showing very different frequency profiles (see Öksüz, 2019 for a discussion). Moreover, while word order is flexible in Turkish, the verb is most typically placed after its object, rather than before it as in English. In contrast, other combination types, such as premodifying adjective + noun combinations, have the same canonical word ordering as English. The priming paradigm relies on presenting a prime word followed by a target word. Thus, in order to study cross-linguistic priming, the normal order in which verbs and their object nouns are presented will need to be violated for either English for Turkish (i.e. to study L1>L2 priming, we must either first show a Turkish verb followed by an English noun (violating Turkish order) or else show a Turkish noun followed by an English verb (violating English word order).

With this range of possible confounding variables in mind, our second research question is:

2. To what extent is cross-linguistic collocational priming affected by
   (a) syntactic structure (premodifying adjective + noun vs. verb + direct object);
   (b) L1-L2 congruence;
   (c) presentation direction (L1-L2 vs. L2-L1);
   (d) collocational frequency.

We assume that adjective + noun collocations are likely to be processed faster cross-linguistically than verb + direct object collocations because the word order of the latter differs between Turkish (i.e. noun + verb) and English (verb + noun). In addition, as previous research has shown that congruent items tend to be more easily processed (Wolter and Gyllstad, 2011; 2013), we may expect to find that items that are congruent between English and Turkish show a stronger priming effect than those that are incongruent. Furthermore, as evidence for priming asymmetry indicates that spreading activation is stronger from the native language to the second language (e.g., Jiang and Forster, 2001), we are more likely to observe a priming effect when the items are presented in the L1-L2 direction rather than the opposite. Finally, we expect to find a negative association between speed of recognition and collocational frequency, in line with earlier research (e.g., Öksüz et al., 2020). In other words, we expect that more frequent collocations will be more strongly entrenched in the mental lexicon, and thus recognized faster.

2. METHODS

2.1 Participants

The participants were preparatory level students of English Language and Literature department at Ankara University. These students can be expected to have a good level of English as all had passed a national foreign language test before being admitted to their degree programmes. In preparation for their intended degrees, these students had also taken English medium classes and intensive courses in receptive and productive language skills during the third and fourth years of their high school education. This is in addition to the usual English language education experienced by students in Turkish high schools. Table 1 shows participants' mean vocabulary size test scores, along with their age, gender and language learning backgrounds (the instruments used to collect these data are described in Section 2.4).

Table 1. Participant Details (N=31; female=20; male=11)

	AGE	VST	Self-reported L2 Proficiency	L2 Immersion	L2 Dominance	L2-L1 Dominance
MEAN	18.79	8,200	0.65	0.43	0.37	0.76
Standard Dev.	2.22	0.70	0.16	0.12	0.10	0.17

VST range: 7,050-10,400
Proficiency, immersion and dominance range 0-1.

2.2 Lexical Decision Task: Procedure

The central tool in this study was a masked primed Lexical Decision Task, created using DMDX[1] (Forster and Forster, 2003). Participants were shown a series of items (described below) on a computer screen. As Table 2 summarizes, each item comprised a focus point (*), which appeared for 500ms, followed by a masking pattern (#########) for 200ms, a *prime* string for 60ms and, finally, a *target* string, which stayed on the screen until the participants responded. Once the target string appeared, participants needed to decide, as quickly as possible, whether the string was a word or a non-word and to record their decision using the computer's keyboard. Their reaction time and the accuracy of their decision was then recorded by the computer. We decided to use a prime word duration of 60ms based on the recommendations of Jiang (2012). As he states, masked priming paradigm helps minimize strategic priming effects and a prime word duration of 50-60ms tends to prevent subjects from detecting the relationship between the prime and target words.

The hypothesis of the experiment is that target items which are preceded by a related prime will be recognised more accurately than those which are preceded by an irrelevant prime. Since the prime item is shown for a very short time, and preceded by a masking pattern, it is expected that primes will not be consciously registered by participants (Jiang, 2012), and that any priming effects detected can therefore be assumed to be at an automatic, rather than conscious, cognitive level.

Table 2. Part of a Sample Priming Experiment Screen in L1-L2

SCREEN 1	SCREEN 2	SCREEN 3	SCREEN 4
* 500 ms	######### 200 ms	prime word 60 ms	target word response is recorded

[1] DMDX is a software designed for psycholinguistic research investigating response times. It was developed at Monash University and at the University of Arizona by K. I. Forster and J. C. Forster and provided as an open-source tool (Forster and Forster 2003). See http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm to download packages and learn more.

*	#########	give (vermek)	priority	Coll
*	#########	take [almak]	priority	Non-coll

As apparatus, we used the computer facilities of Ankara University, School of Foreign Languages (20 computers with exactly the same capabilities). The computers (HP Compaq 8200 Elite SFF) have Intel İ5 -2400 CPU, 3.10 GHz processors and 4096 MB rams. Their display mode is 1600x900 (32 bit), (75Hz). We used standard English keyboards and recorded reaction time through key strokes (Right-Left CTRL).

2.3 Lexical Decision Task: Items

Item lists were developed separately for English and Turkish. In both cases, the lists comprised two types of noun-centred collocation: ADJ+N collocations and V+N collocations. In each case, nouns were intended as target words and adjectives/verbs as primes. For the English list, the Corpus of Contemporary American English (COCA) list, created by Davies (2008–) was used as a starting point. This list contains node words and collocate pairs for the top 5,000 headwords from COCA; approximately 732,000 collocates. The file contains those collocates that occur at least three times with the node word and which have a Mutual Information (MI) score of 1.0 or higher.

Candidate items were identified as ADJ+N or V+N combinations meeting the following criteria, which were applied in turn:

- node noun is amongst the 3,000 most frequent nouns on the list
- collocation has an MI score of at least 3.0 and *t*-score of at least 2.0
- collocation has a two-word counterpart in Turkish with no case marking for the nouns and no tense marking or person suffixes for the verbs (i.e. without any inflectional affixes)
- component words do not have cognates in Turkish (e.g. collocations including *music* and *restaurant* were excluded due to their resemblance to their Turkish equivalents *müzik* and *restoran*).

The first author of this article next created a list of Turkish counterparts to these items such that each collocation represented a correct translation of the English collocation. Association measures for these Turkish items were calculated using data from the Turkish National Corpus (TNC) (Aksan et al., 2012). Where items failed to reach an MI of 3.0 or *t*-score of 2.0 in this corpus, both English and Turkish versions were deleted from the lists.

Collocations meeting these criteria were further filtered such that only one item was included for each node noun. In selecting items, the first author attempted to extract an even mix of congruent and incongruent collocations, based on his subjective judgement as an L1 Turkish and advanced L2 English user. "Congruence" of collocations was defined as the equivalence of word combinations in terms of their meaning in L1 and L2. As described in the review above, the effects of congruence on collocational processing have been discussed in earlier research (e.g., Wolter and Gyllstad, 2013). To confirm the final categorization of items as congruent or non-congruent, we asked four L1 Turkish-speaking instructors of English language who have been working full time at a Turkish state university to help us with the filtering. We explained to them the concept of "congruence" for collocations, as defined above, and gave them some examples from Turkish and English. We then gave them a list of our target collocations in English and asked them to tag the items as "congruent", "non-congruent" or "not sure" considering their Turkish counterparts. Those items which were tagged the same (either as congruent or incongruent) by at least three experts were kept on the list (*n*=37 V+N and 40 ADJ+N).

As the last step, we consulted the website, English Vocabulary Profile[2], to check the CEFR (Common European Framework) levels of each item (Council of Europe, 1996). Seven items which were C1-C2 level according to the CEFR were eliminated from the final list. All remaining items were at A2 to B2 level. Our decision for the CEFR level can be attributed to the vocabulary size test scores of our participants (Mean=8,200). Nation and Beglar (2007) state that competent L2 users who are approximately at C1-C2 level according to CEFR tend to have approximately a 9,000 word vocabulary. Based on their remarks, we decided to exclude C1-C2 level words as we believe, those possibly unknown words could distort the results. From the final list of items (*n*=70), we chose 60, distributed equally across part-of-speech and congruence categories, as summarized in Table 3.

Table 3. Item Dispersion by Part of Speech (POS) and Congruence

Experimental Items			
Congruent		Non-congruent	
V+N	ADJ+N	V+N	ADJ+N
15 coll.	15 coll.	15 coll.	15 coll.
*See Appendix A for the list of target collocations in the study			

Another important issue for the priming experiment was the production of non-collocate items, fillers and non-words. Non-collocate items paired target nouns from the experimental materials above with adjective/verb primes with which they do not commonly collocate. To qualify as non-collocations, and based on the suggestions of Schmitt (2010) and Stubbs (1995) the prime + target combinations needed to occur with *t*-scores of lower than 1.0 and MI scores

[2]See the website http://vocabulary.englishprofile.org/staticfiles/about.html for more details.

of lower than 2.0 in COCA or the TNC, depending on the language. (Mean= -0.02 for *t*-score; 0.3 for MI in Turkish and -58.9; -7.75 in English, respectively). To neutralize the effect of first and second exposure to the same noun on the RTs across the control or experimental conditions and to ensure that participants were not able to recognize any discernible patterns in the priming experiment, the order of items was randomized. To account for possible effects of the frequency of prime words on processing, primes in non-collocate members had similar frequency values with primes in the collocate items. The statistical comparison of the frequency values did not indicate any significant differences across conditions. Table 4 presents a detailed comparison of the log-transformed word level frequency levels of the prime words.

Table 4. Median (IQR) log-transformed Word Level Frequency of Collocate and Non-collocate Primes

	V+N				ADJ+N			
	Congruent		Non-congruent		Congruent		Non-congruent	
	Collocation	Non-collocation	Collocation	Non-collocation	Collocation	Non-collocation	Collocation	Non-collocation
En	4.73 (0.74)	4.78 (0.66)	4.48 (0.73)	4.51 (0.77)	4.52 (0.57)	4.53 (0.67)	4.56 (0.61)	4.52 (0.67)
Tk	1.82 (0.69)	1.61 (0.43)	2.15 (0.86)	2.08 (1.15)	2.11 (0.61)	2.10 (0.63)	1.78 (1.56)	1.76 (1.31)

Filler items were non-collocate random word combinations which included lexical items (both prime and target words) consisting of same (-/+1) number of letters as experimental items to control for the word-length effect. The frequency values of the lexical items used as fillers were not controlled as the reaction times of those items were not included in the final analysis. Non-words were chosen from the ARC non-word database[3] (Rastle et al., 2002) for the English words and the Wuggy Pseudo word generator (Keuleers and Brysbaert, 2010) for the Turkish items.

Finally, the two sets of monolingual items described above were combined to create cross-linguistic materials for the experiment. Each item was assigned to either an L1-L2 (Turkish prime-English target) or an L2-L1 (English prime-Turkish target) condition. Assignment of items to conditions was randomized, with the condition that, within each part-of-speech x congruence condition, approximately equal numbers were assigned to each condition, as illustrated in Table 5.

Table 5. Item Dispersion of Collocations

Congruent		Non-congruent	
V+N	ADJ+N	V+N	ADJ+N

[3] The website can be accessed at http://www.cogsci.mq.edu.au/research/resources/nwdb/nwdb.html.

(7/8 coll.) L1-L2 / L2-L1	(7/8 coll.) L1-L2 / L2-L1	(7/8 coll.) L1-L2 / L2-L1	(7/8 coll.) L1-L2 / L2-L1
*60 word combinations (collocations only) *120 word combinations (collocate/non-collocate items) *260 (collocate/non-collocate/non-word/filler items)			

2.4 Instruments for Additional Data

Language History Questionnaire 3.0[4]

We conducted a language background questionnaire developed at Penn State (Li et al., 2006) as a standard procedure to report our participants' details. There were sections about age, gender, educational background, language exposure and immersion, and self-rated proficiency.

Vocabulary Size Test (VST)[5]

We used the Vocabulary Size Test by Nation and Beglar (2007). We assumed that since the priming experiments consisted of word patterns (i.e. collocations), assessing participants' vocabulary level, rather than their overall proficiency, would fit better with the overall aims of our study. Some research studies also indicate that there is a positive correlation between proficiency level and vocabulary size test scores. For example, Stæhr (2008) claims based on a national proficiency test in Denmark that learners' vocabulary size correlates significantly with their reading and writing abilities. In addition, Miralpeix and Muñoz (2018) state that vocabulary size can be a good indication of language proficiency, for students with a vocabulary size of over 5,000 word families in particular. Milton (2010) states that there is a relationship between receptive vocabulary size scores and CEFR levels. Finally, Meara (1996) asserts that L2 users with larger vocabulary sizes are likely to be more proficient in a wide range of language skills as opposed to L2 users with smaller vocabulary sizes. He further states that vocabulary skills have potential to contribute to all aspects of L2 proficiency. Therefore, the possible comparison between the vocabulary size and language proficiency in our study is borrowed from Milton and Meara's arguments.

Post-Experiment Questionnaire

Participants were asked to answer questions regarding dexterity and vision. None of the participants reported any serious problems with their eyesight that could hinder their performance in the experiment. The majority of the participants reported they were right-hand dominant. One participant stated he used both his hands frequently. The responses of all the participants who took the experiment were accepted before data trimming.

[4] The questionnaire can be accessed at http://blclab.org/lhq3/.
[5] See https://my.vocabularysize.com for more information.

2.5 Data Analysis

Based on the recommendations of Jiang (2012), the responses of one participant, who had an error rate of more than 20%, were eliminated from the analysis. This resulted in one participant's data being removed. Additionally, all responses faster than 200 milliseconds or slower than 2000 milliseconds were trimmed from the final data. As Jiang (2012) states, reaction times lower than 200 milliseconds should be considered unacceptable since it is unlikely that they result from genuine word recognition process. Additionally, if the reaction time is slower than 2000 milliseconds, it is very likely that the participant is either using a strategy (thus the processing is not automatic) or is distracted by a spurious item. 93.4% of the response times (RTs) were kept for the statistical analysis. As Jiang (2012) explains, data treatment can affect 3-7% of the data. If more than 10% of the data are lost after the treatment, the criteria for data treatment need to be reconsidered.

In our statistical models, the outcome variable was log-transformed reaction time (reaction time was log-transformed to reduce the skewness in the distributions, as suggested by Siyanova et al., 2011). The predictor variables were: presentation direction (L1-L2 vs. L2-L1); part of speech (POS); congruence; collocation frequency (*t*-score), collocation MI (though this was removed later due to multicollinearity concerns) and collocation Delta-P. Delta-P is a measure of collocational association which, unlike MI, is directional: that is, it distinguishes cases in which the first word predicts the second word (e.g. *upside down*) from cases where the second word predicts the first (e.g. *of course*). It provides two separate statistics for each direction (Gries, 2013). To avoid confusion in the current article due to the variable word order of Turkish, we refer to the statistics as:

- Delta-$P_{verb|noun}$: showing how strongly the verb is predicted when the noun occurs
- Delta-$P_{noun|verb}$: showing how strongly the noun is predicted when the verb occurs
- Delta-$P_{adjective|noun}$: showing how strongly the adjective is predicted when the noun occurs
- Delta-$P_{noun|adjective}$: showing how strongly the noun is predicted when the adjective occurs

Additionally, when we report the statistics with the items in the two structures merged (i.e. ADJ+N and V+N together), we refer to them as Delta-P word 1 > word 2, Delta-P word 2 > word 1.

All statistical calculations in this article were performed using jamovi version 0.9 (2019) and GAMLj (General Analyses for Linear Models) by Gallucci (2019). See Appendix B for the model summary to run the same analysis on R. To determine the relationship between predictor variables and reaction times, we conducted a mixed effects model analysis (Baayen et al., 2008). A mixed effects model is preferred over a regression model in psycholinguistic research because if a subjects-analysis (averaging over items) is carried out, the by-item variation is essentially disregarded. Conversely, in the items-analysis, by-subject variation is ignored and mixed models account for both sources of variation in a single model (Winter, 2019). ‘Item’ and ‘participant’ variables were entered as crossed random effects to take into account item-specific idiosyncrasies and idiosyncratic variation due to differences between participants. All other predictor variables (Delta-P word 1 > word 2, Delta-P word 2 > word 1 for both structures; *t*-score in English and Turkish) were entered as covariates. The model fitting procedure started with a maximal model that included log-transformed reaction time as the outcome variable and all potential predictor variables as main effects. The best fitting model was found by eliminating, one by one, the predictor variable that had the least impact on the model outcome until only the variables that significantly improved the fit were included. The fixed effects; congruence, POS, the direction of language presentation and their interactions were kept in each analysis on purpose although they did not reveal any significant associations in certain occasions in an attempt to illustrate the same effects in each output. Deviation contrast coding was used for the fixed effects as this provided us with some interpretational advantages for analyzing the interactions.

Before building a mixed effect model, possible correlations between the independent variables indicating collocational frequency were explored to avoid multicollinearity (i.e. strong inter-associations among the independent variables). Myers (1990) suggests that a variance inflation factor (VIF) of 10 is a good value at which to worry. Additionally, tolerance values below 0.1 indicate serious problems although Menard (1995) suggests that values below 0.2 are worthy of concern. In our study, MI scores in both Turkish and English had a VIF of 20.29 and 24.29, respectively and their tolerance values were lower than 0.1. We therefore decided to keep Delta-P and *t*-score in both languages for our final analysis and eliminate the association measures with stronger correlations. The VIF and tolerance scores of each association measure can be seen in Table 6. Finally, the covariates in our analysis were centered as suggested by Field (2016) to help combat multicollinearity between predictor variables.

Table 6. Collinearity Statistics

	VIF	**Tolerance**
TR *t*-score	3.07	0.3258
TR MI	20.29	0.0493
TR Delta-P **word 1 > word** 2 (for both structures)	1.49	0.6717
TR Delta-P **word 2 > word 1** (for both structures)	2.20	0.4538
ENG *t*-score	1.84	0.5437
ENG MI	24.79	0.0403
ENG Delta-P **word 1 > word 2** (for both structures)	1.68	0.5969
ENG Delta-P **word 2 > word 1** (for both structures)	1.37	0.7295

3. FINDINGS

This section reports on the associations between RTs (of both the collocate and non-collocate items) and the independent variables (fixed effects: the direction of language presentation, POS, congruence, and collocational frequency values in both languages). Table 7 presents the results of the best-fitting mixed effects model for our experimental items.

Table 7. Mixed Effects Model (All Items)

Effect	Estimate	SE	df	t	p
(Intercept)	2.781	0.014	25.8	195.30	< .001
L1-L2 vs. L2-L1	0.016	0.042	51.7	3.914	< .001
POS	-0.070	0.0043	51.7	-1.606	.114
Congruence	-0.0017	0.0041	51.7	-0.413	.681
Delta-P $_{\text{word 1 > word 2}}$ (English)	-0.232	0.081	51.7	-2.582	.006
t-score (Turkish)	0.0016	9.22e-4	52.0	1.77	.08
L1-L2 vs. L2-L1 * POS	-0.010	0.0043	51.4	-2.342	.02
Congruence * L1-L2 vs. L2-L1	-0.0089	0.0043	51.7	-2.048	.04
Random Effects	**Variance**				**SD**
Item	4.46e-4				0.021
Subject	0.004				0.068
Goodness of Fit	R^2 conditional .28 / R^2 marginal .03				

This model demonstrates two main effects: direction of language presentation (targets in the first language were responded to faster than targets in the second language) and part-of-speech (nouns following adjective primes were processed faster than those following verb primes) when in interaction with the direction of language presentation. Although the participants naturally processed L1 words faster, particularly when the ADJ+N items were presented in L1-L2 direction, collocate items were processed faster than non-collocate items. To be more precise, the processing durations of the ADJ+N collocations in L1-L2 direction were potential indicators of cross-linguistic collocational priming. The only significant collocational predictor was Delta-P word 1 > word 2 in English. Since the numbers indicate a negative association, it can be claimed that a higher Delta-P in English leads to lower RTs, which can tentatively be regarded as evidence for cross-linguistic collocational priming. Another noteworthy finding was the effect of congruence and presentation direction on the RTs. Those collocations which are both congruent between the first and second language and presented in L1-L2 direction were associated with faster RTs. The same effect could not be seen for the congruence variable alone.

Since part-of-speech clearly has a strong effect on reaction times when in interaction with the direction of language presentation and assuming that the different word order in the two languages has an influence on processing, we split our data into separate sets for ADJ+N and V+N combinations and re-ran the mixed effects models separately on each data set. The best-fitting models for these analyses are shown in Tables 8 and 9.

Table 8. Mixed Effects Model (ADJ+N)

Effect	Estimate	SE	df	t	p
(Intercept)	2.772	0.014	24.6	194.89	< .001
L1-L2 vs. L2-L1	0.0056	0.0051	22.1	1.090	.28
Congruence	0.0045	0.0050	22.2	0.906	.37
Delta-$P_{adj.\vert noun}$ (English)	-0.224	0.077	22.1	-2.889	.008
t-score (English)	-6.22	3.51e-4	22.4	-1.771	.09
t-score (Turkish)	0.0029	0.0014	22.8	2.002	.05
L1-L2 vs. L2-L1 * Congruence	-0.0169	0.0048	22.3	-3.328	.003
Random Effects		**Variance**			**SD**
Item		1.14e-4			0.010
Subject		0.004			0.067
Goodness of Fit		R^2 conditional .28 / R^2 marginal .03			

As Table 8 shows, adjective + noun combinations followed a similar trend to that seen in Table 7 for the full data-set. That is, although reaction times are not associated with direction of language presentation, they seem to be associated with the variables of congruence and the direction of language presentation when they are in interaction. More specifically, when the ADJ+N items were presented in L1-L2 and they were congruent across L1 and L2, participants responded to the target nouns faster than the control items. Additionally, the association measures, Delta-P in English and *t*-score in English [near statistical significance], had negative associations with the RTs. That is, higher Delta-P and *t*-scores led to faster RTs and thus we can claim that the target nouns were primed in the L1 Turkish-L2 English participants' mental lexicon.

Table 9. Mixed Effects Model (V+N)

Effect	Estimate	S	df	t	p
(Intercept)	2.791	0.0153	26.8	182.44	< .001
L1-L2 vs. L2-L1	0.044	0.0123	24.6	3.662	.001
Congruence	-0.004	0.0012	24.6	-0.334	.74
Delta-$P_{verb\vert noun}$ (Turkish)	0.039	0.0138	24.3	2.870	.008
L1-L2 vs. L2-L1 * Congruence	0.030	0.0061	24.6	0.505	.61
Random Effects		**Variance**			**SD**
Item		4.82e-4			0.021
Subject		0.004			0.072
Goodness of Fit		R^2 conditional .29 / R^2 marginal .04			

On the other hand, as far as verb + noun collocations are concerned, those combinations are in association both with direction of language presentation and with a different collocational variable: the strength of prediction from noun to verb in Turkish, as measured by Delta-$P_{verb\vert noun}$. Interestingly, a stronger collocational association was associated with slower reaction times. The congruent items were not processed faster than the non-congruent items and, in contrast to the analysis for ADJ+N combinations, their interaction with the direction of language presentation did not make any difference. Lastly, when items were presented in the L2-L1 direction (that is, when the verb prime was in English and the noun target was in Turkish), the participants responded to the target items faster.

These models point to a number of noteworthy effects. As for the ADJ+N collocations, disregarding the direction of presentation, Delta-P in English results in faster RTs in the cross-linguistic priming experiment. We found evidence of a congruence effect particularly for those items in the L1-L2 direction, which is line with earlier research (e.g. Wolter and Gyllstad, 2013). The evidence suggests that the same word order for ADJ+N collocations in Turkish and English appears to facilitate cross-linguistic collocational processing and the fact that Delta-P and *t*-scores are significant predictors of RT can be regarded as tentative evidence for a collocational priming effect. In addition, as earlier research (Wolter and Gyllstad, 2011) also states, spreading activation is more robust in the L1-L2 direction when the target lexical items semantically overlap in the L1 and L2.

On the whole, it can be claimed that an ADJ+N combination is more likely to trigger collocational priming in the L1 Turkish-L2 English bilingual mental lexicon and that the collocational frequency values Delta-$P_{adj|noun}$ and *t*-score are better predictors of RT than other investigated frequency values. Additionally, although we know that participants naturally respond faster when items are presented in the L2-L1 direction, possible cross-linguistic collocational interaction is more robust when they are presented in the L1-L2 direction. This is shown by the mean reaction time differences of the collocate and non-collocate items in that direction and the detected interactions. Finally, the dominant language of the bilinguals might be causing an inhibition effect as we have found that Delta-$P_{verb|noun}$ in Turkish has a positive correlation with reaction times.

4. DISCUSSION

According to the network model of language processing (Fortescue, 2014), when a language user hears or produces a lexical item (e.g. *heavy*) immediately followed by another lexical item (e.g. *rain*), a strong link is formed between them. Depending on how frequently this combination is encountered, when the language user hears or produces the node *heavy*, the collocate *rain* is mentally activated—together with other semantically related items the person has been exposed to in his/her language learning experience. As Anderson (2005: p. 455) defines it, mental activation is “a state of memory traces that determines both the speed and the probability of access to a memory trace”. Hoey’s (2005) theory of *Lexical Priming*, which is also the starting point of our research, bears great resemblance to the network model of language processing in many aspects. He claims that any word which tends to co-occur frequently with another word is mentally activated when the first word (or the second word, depending on bidirectionality) is uttered or heard. As also asserted by the collocational spreading activation model (Wolter and Yamashita, 2014), repeated exposure to frequently co-occurring patterns makes the associations between words stronger and more robust. Research also suggests that the frequency of word combinations or multiword units appear to influence the processing of those lexical patterns (Jolsvai, McCauley and Christiansen, 2013). The traces of the network model can also be found in Construction Grammar. As Langacker (1987) states, constructions exist in a network and when we hear or produce a construction, a pattern of nodes is activated in our internal lexicon. Depending on how frequently these nodes are activated, the construction may be entrenched and processed holistically as a single unit. When a construction is entrenched, its activation becomes so highly automated that no conscious attention is needed when processing it (Wolter and Gyllstad, 2013). We have hypothesized that if the mental lexicon is wired to store language in chunks and the activation of a single constituent part spreads to others as is indicated in the theories discussed above, then it is likely that it imitates this mechanism in L2 processing. In other words, if we pay attention to chunks during language production to ease processing burden in the L1, then it is possible that we follow a similar path while learning a second language. As reviewed above, research to date on the L2 processing of collocations suggests that this is indeed the case (Gyllstad and Wolter, 2016; Wolter and Gyllstad, 2011; 2013; Wolter and Yamashita, 2014; Yamashita, 2018).

The present study has attempted to extend this picture by extending it to an understudied language (Turkish) and investigating the extent to which the activation of individual words spreads to the collocates of those words across languages. That is, it examines whether collocational access in the bilingual lexicon is language non-selective (Hermans et al., 1998). While we found no evidence of cross-linguistic collocational priming for V+N combinations, we did identify some evidence of such effects for ADJ+N combinations. Specifically, when a first language (Turkish) prime was seen before a second language (English) target, the speed of recognition of the target was negatively associated (indicating faster RTs) with the combination's Delta-$P_{adj.|noun}$ and *t*-scores. More strikingly, when the ADJ+N collocations were presented in L1-L2 direction and particularly when they were congruent, the participants' recognition speed was increased. We know from earlier research that cross-linguistic overlap tends to trigger greater priming at the lexical level (Wolter and Gyllstad, 2011; 2013). The negative association we detected between collocational frequency and the speed of recognition is also in line with earlier research. Wolter and Gyllstad (2013) and Öksüz et al. (2020) also state that higher collocational frequency in L2 faciliates the processing of collocational items in the mental lexicon, which is likely to result in a greater priming effect.

Unlike ADJ+N collocations, the speed of recognition for V+N items was positively associated (indicating slower RTs) with its Turkish Delta-$P_{verb|noun}$. Additionally, when a second language (English) prime was seen before a first language (Turkish) target, the recognition of the target word was faster and congruency had no role in this process. This finding contradicts Wolter and Gyllstad (2011), who found a priming effect in V+N collocations, particularly when they were congruent between English and Swedish. In each case, the relationship between the fixed predictor variables is rather weak, whereas the associations with the random variables of item and (most strongly) subject are much stronger.

We acknowledge that our findings pose rather more questions than they answer, but our hope is that putting these questions forward will provide a good foundation for further work. The first concerns the overall weak model fits between the fixed effects and reaction times and the apparently strong effects of subjects. This suggests a good deal of diversity within this particular group of participants. An important point to note here is that our participants could be at a lower level of English proficiency than those who took part in previous collocational priming studies with Turkish learners (Cangır, 2018a; Öksüz, 2019). In Cangır (2018a), the participants were experienced instructors of English who worked at state universities in Turkey and Öksüz (2019) had participants who were advanced users of English evidenced by their vocabulary knowledge test scores. It is therefore possible that learners at this level are able to benefit from collocational priming in their L2 only to a limited extent. Although we found evidence for cross-linguistic collocational priming for ADJ+N word combinations, further research investigating more advanced L2 users is necessary to find more robust results. It should be noted that cross-linguistic priming effects more generally are strongest in more advanced learners (Altarriba and Basnight-Brown, 2007). Additionally, according to Van Hell and Tanner (2012), as proficiency in L2 increases and L2 users tend to code-switch between L1 and L2 more often, the cognitive mechanisms influencing the activation and inhibition of the two systems may change. To put it another way, attentional and cognitive control develop thanks to increased proficiency in the L2 and advanced bilinguals have a stronger ability to filter out irrelevant information which could interfere with lexical processing and they enjoy a processing advantage (Segalowitz and Hulstijn, 2005; Bialystok and Craik, 2010). It would therefore be useful to replicate the studies conducted here with a more advanced group of learners. Additionally, since our participants switch between seeing the two directions, L1-L2 and L2-L1, during the study, it may be the code-switching effect which shadows some of the findings due to a processing burden in their mental lexicon. Evidence from earlier research suggests that code-switching has cognitive costs (van Heuven, Dijkstra and Grainger, 1998). Thus, an acceptability judgements task with L2 only items could yield more reliable results and future research may need to look at the issue from a different angle to rule out a possible cognitive burden effect.

A second issue arising from our findings concerns the associations between Delta-$P_{verb|noun}$ seen in the Turkish verb prime + English noun target condition and the speed of recognition. The positive correlation between Delta-$P_{verb|noun}$ and reaction times points to potential complexity: when the noun target is a strong predictor of the verb prime, recognition is slowed. We can speculate that a process of inhibition may be be occurring here, whereby re-activation of the recently-seen prime leads to hesitation in registering recognition of the noun.

Given the cross-linguistic difference between English and Turkish (with the latter favouring pre-verbal objects), V+N collocations can be predicted to show less priming. In addition, we can tentatively claim that a strong directional collocational association in Turkish may have induced interference with the English equivalent. However, understanding this effect fully will require further investigation.

The finding that Turkish adjective primes facilitated recognition of English noun targets in cases where the collocations were congruent between the two languages is in line with previous research on collocation priming within a second language (Wolter and Gyllstad, 2011; 2013; Wolter and Yamashita, 2014; 2017) and also suggests a cross-linguistic priming effect. Earlier research (e.g., Jiang and Forster, 2001; Zhao et al., 2011) has suggested that there is priming asymmetry in cross-linguistic lexical interaction in the bilingual mental lexicon in that first language words facilitate second language words more strongly than vice-versa, so it is to be expected that effect will differ across the two directions. However, the reasons for the specific differences found are not yet clear.

Considering our tentative findings, prior assumptions and the related research in the literature, we believe one of the best fitting models to explain collocational priming and its effect on the structuring of mental lexicon is the *Spreading Activation Model* (Collins and Loftus, 1975). This model has also been addressed in recent research studies exploring collocational priming in L2 (e.g., Wolter and Gyllstad, 2013). Wolter and Gyllstad (2013) state that like lexical items with semantic associations, collocational items in L1 and L2 are linked to one another at different strengths depending on how entrenched they are in the lexicon. Considering the inhibition effect suggested by the Delta-P value in our priming experiment, another significant model can be regarded as the *BIA* (Bilingual Interactive Activation) *Model* (Dijkstra and Van Heuven, 1998). The model, embracing a language non-selective lexical access approach, emphasizes the robust interconnectedness between lexical items not only within but also across languages. The interconnectedness is influenced by certain factors (e.g. frequency) resulting in an inhibitory or a facilitative effect on the processing of lexical items (See Cangır, 2018b for a brief review of mental lexicon models). In an attempt to crudely illustrate the possible flow of the cross-linguistic activation at the collocational level, particularly for congruent adjective-noun word combinations in Turkish and English, we present Figure 1, which was adapted from Conklin and Carrol (2018).

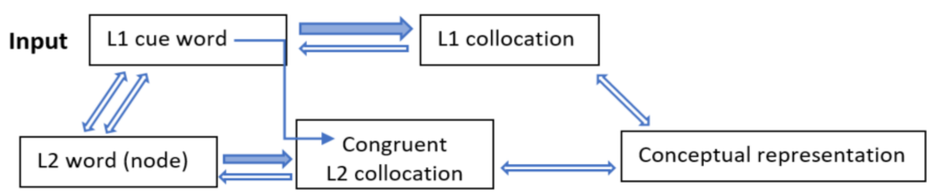

Figure 1. Collocational Spreading Activation in L1-L2 direction (Congruent items)

In addition to the evidence provided by prior priming experiments, there is also research indicating that collocations are processed faster than non-collocations through an EEG (Electroencephalogram) experiment (Hughes and Hardie, 2019). So we have neurological evidence of psycholinguistic reality of collocations though not at the cross-linguistic level. We can assume then that if the constituent parts of a collocation are linked to one another at the neurological level in L1, then there can be links at the L1-L2 level assuming that the lexical items in the two lexicons interact with each other during language processing. As Hughes and Hardie (2019) state, the lexical items (L1 and L2) in the bilingual mental lexicon form links (semantic, collocational etc.) with varying strength, which may be considered similar to the pattern of interconnected neurons in the brain. However, it must be noted that the existence of cross-linguistic collocational spreading activation at neurological level is yet to be addressed and validated in further research.

Our results are partly in line with earlier research (e.g. Cangır et al., 2017; Cangır, 2018a), but also tentatively elaborate their findings. In addition, we extend Hoey's (2005) idea of collocational priming by investigating the issue from a cross-linguistic perspective focusing on an understudied and morphologically rich language with largely non Indo-European vocabulary. Future research is needed to find stronger evidence for cross-linguistic spreading activation and collocational priming in L2 by more controlled experimental settings and using more sophisticated tools, such as EEG or eye-tracking (Siyanova-Chanturia, 2013).

5. CONCLUSION

This study found evidence for cross-linguistic collocational priming for ADJ+N collocations in particular. It was also asserted that those word combinations were processed faster when they were presented in the L1-L2 direction and when they were congruent between the two investigated languages. These findings support the language non-selective lexical access at the collocational level and indicate that L1 Turkish-L2 English users might be sharing a single lexicon with items from either language or separate lexicons with interconnected lexical items with varying strength. As for the V+N collocations, as the positive association between Delta-P and reaction times indicate, an inhibition effect was in place, possibly due to the different word order of the two languages (i.e. V+N in English vs. N+V in Turkish), which seemed to result in a processing burden in the participants' mental lexicon. There is need for more controlled research scrutinizing both the variables in this study and different aspects, such as learner differences, proficiency levels etc. to reach stronger conclusions regarding collocational processing in the L1 Turkish-L2 English mental lexicon. We hope the methodology and the results of this study may trigger further research investigating cross-linguistic lexical processing in the Turkish context and stimulate studies focusing on phraseology from the eyes of understudied languages like Turkish.

APPENDIX A

List of Items (English)

VERB +NOUN		ADJ+NOUN	
Congruent		***Congruent***	
1	make a mistake	1	deep sleep
2	give permission	2	cold war
3	take pleasure	3	outside world
4	find solace	4	strong evidence
5	show affection	5	naked eye
6	take a breath	6	warm welcome
7	find a solution	7	bitter end
8	commit murder	8	heated debate
9	give priority	9	rich history
10	make a discovery	10	golden age
11	find a clue	11	middle class
12	break heart	12	opposing view
13	open fire	13	high court
14	win a victory	14	undying love
15	pass time	15	white lie
Non-congruent		***Non-congruent***	
1	make a decision	1	open mind
2	pay attention	2	long run
3	cast doubt	3	heavy rain
4	go bankrupt	4	thick smoke
5	take a break	5	wiry hair
6	feel the need	6	strong smell
7	put pressure	7	false tooth
8	lose weight	8	strong coffee
9	pay a visit	9	soft drink
10	shed light	10	driving force
11	set an example	11	tall building
12	grow beard	12	high achievement
13	have an accident	13	sharp fall
14	place emphasis	14	drastic change
15	keep a secret	15	free rein

List of Items (Turkish)

NOUN+VERB		ADJ+NOUN	
Congruent		***Congruent***	
1	hata yap-	1	derin uyku
2	izin ver-	2	soğuk savaş
3	keyif al-	3	dış dünya
4	huzur bul-	4	kuvvetli delil
5	şefkat göster-	5	çıplak göz
6	nefes al-	6	sıcak karşılama
7	çözüm bul-	7	acı son
8	cinayet isle-	8	ateşli tartışma
9	öncelik ver-	9	zengin tarih
10	keşif yap-	10	altın çağ
11	ipucu bul-	11	orta sınıf
12	kalp kır-	12	karşıt görüş
13	ateş aç-	13	yüksek mahkeme
14	zafer kazan-	14	ölümsüz aşk
15	zaman geçir-	15	beyaz yalan
Non-congruent		***Non-congruent***	
1	karar ver-	1	açık fikir
2	dikkat et-	2	uzun vade
3	şüphe uyandır-	3	sağanak yağmur
4	iflas et-	4	yoğun duman
5	ara ver-	5	kabarık saç
6	ihtiyaç duy-	6	keskin koku
7	baskı yap-	7	takma diş
8	kilo ver-	8	koyu kahve
9	ziyaret et-	9	alkolsüz içki
10	ışık tut-	10	itici güç
11	örnek ol-	11	yüksek bina
12	sakal bırak-	12	büyük başarı
13	kaza yap-	13	sert düşüş
14	vurgu yap-	14	köklü değişiklik
15	sır sakla-	15	tam yetki

APPENDIX B

Model Info

Info	
Estimate	Linear mixed model fit by REML
Call	Log_RT ~ 1 + Language-n + POS + congruence-n + TR_t-score + ENG_Delta P_1 + Language-n:POS + Language-n:congruence-n+(1 | Item)+(1 | Subject)
AIC	-1944.8921
BIC	-1818.3835
R-squared Marginal	0.0314
R-squared Conditional	0.2835

WORKS CITED

Aksan, Y. et al. (2012). Construction of the Turkish National Corpus (TNC). In Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC 2012). İstanbul. Turkiye. http://www.lrec-conf.org/proceedings/lrec2012/papers.html

Altarriba, J. (1992). The representation of translation equivalents in bilingual memory. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 157-174). Amsterdam: Elsevier.

Altarriba, J., & Basnight-Brown, D.M. (2007). Methodological considerations in performing semantic- and translation-priming experiments across languages. *Behavior Research Methods*, *39*(1), 1–18. https://doi.org/10.3758/BF03192839

Anderson, J. R. (2005). *Cognitive psychology and its implications* (6th ed.). New York, NY: Worth Publishers.

Baayen, R.H., Davidson, J. D., & Douglas, M.B. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: the role of language direction and language dominance. *Memory & Cognition*, *35*(5), 953–965. https://doi.org/10.3758/BF03193468

Bialystok, E., & Craik, F.I.M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions Psychological Science*, *19*(1), 19–23.

Cangır, H., Büyükkantarcıoğlu, S.N., & Durrant, P. (2017). Investigating Collocational Priming in Turkish. *Journal of Language and Linguistic Studies*, 13(132), 465–486. Retrieved from www.jlls.org

Cangır, H. (2018a). *Investigating the Relationship between L1 and L2 Collocational Processing in the Bilingual Mental Lexicon* (Doctoral dissertation, Hacettepe University). Retrieved from http://www.openaccess.hacettepe.edu.tr:8080/xmlui/handle/11655/4986

Cangır, H. (2018b). Bilingual Mental Lexicon and Collocational Processing. In D. Buğa, & M. Coşgun Ögeyik (Eds.), *Psycholinguistics and Cognition in Language Processing* (pp. 221-243). Hershey, PA: IGI Global.

Carrol, G., & Conklin, K. (2015). Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs.' *Bilingualism: Language and Cognition*, *20*(2), 1–19. https://doi.org/10.1017/S1366728915000103

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428. doi:10.1037/0033-295X.82.6.407

Conklin, K., & Carrol, G. (2018). First language influence on the processing of formulaic language in a second language. In A. Siyanova-Chanturia, & A. Pellicer-Sanchez (Eds.), *Understanding Formulaic Language. A Second Language Acquisition Perspective.* (pp. 62-77). Routledge.

Council of Europe (1996). *The Common European framework of reference for languages: Learning, teaching and assessment*. http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf. Accessed 1 February 2021.

Davies, M. (2008-). The Corpus of Contemporary American English: 520 million words, 1990-present. Available online at http://corpus.byu.edu/coca/.

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Scientific psychology series. Localist connectionist approaches to human cognition* (p. 189–225). Lawrence Erlbaum Associates Publishers.

Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, 6(2), 125–155.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, 26(2), 163–188.

Durrant, P. (2013). Formulaic morphology in an agglutinating language: the case of Turkish.*Corpus Linguistics and Linguistic Theory,*9(1): 1-38.

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. (2008). The frequency effect in second-language visual word recognition: evidence against a language-selective rank hypothesis. *Psychonomic Bulletin and Review*, 15(4), 850– 855.

Field, A. (2016). *An adventure in statistics: the reality enigma.* Quantitative/Statistical Research. SAGE Publications, London. ISBN 9781446210451

Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1), 116–124. http://doi.org/10.3758/BF03195503

Fortescue, M. (2014). A network model of lexical organisation in the bilingual brain. *4th UK Cognitive Linguistics Conference*, 117–135.

Gallucci, M. (2019). GAMLj: General analyses for linear models. [jamovi module]. Retrieved from https://gamlj.github.io/.

Gries, S.T. (2013). 50-something years of work on collocations: What is or should be next ... *International Journal of Corpus Linguistics*, 18(1), 137–166.

Gyllstad, H., & Wolter, B. (2016). Collocational processing in the light of a phraseological continuum model: Does semantic transparency matter? *Language Learning*. 66(2), 296-323

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing word in a foreign language: Can speakers prevent interference from their first language? Bilingualism: *Language and Cognition*, 1(3), 213–230. doi:10.1017/ S1366728998000364

Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. Routledge.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19(1), 24–44.

Huang, L.S. (2001). 'Knowledge of English Collocations: An Analysis of Taiwanese EFL Learners', in C. Luke and R. Rubrecht (eds), *Texas Papers in Foreign Language Education: Selected Proceedings from the Texas Foreign Language Education Conference 2001*. Educational Resources Information Center Document ED465 228.

Hughes, J., & Hardie, A. (2019). Corpus Linguistics and Event-Related Potentials. In J. Egbert, & P. Baker (Eds.), *Using corpus methods to triangulate linguistic analysis* (pp. 185-218). London: Routledge.

Jiang, N., & Forster, K. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language*, 44, 32-51.

Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York: Routledge.

Jiang, N. (2015). Six decades of research on lexical representationand processing in bilinguals. In J. Schwieter (Ed.), *The Cambridge Handbook of Bilingual Processing* (Cambridge Handbooks in Language and Linguistics, pp. 29-84). Cambridge: Cambridge University Press. doi:10.1017/CBO9781107447257.002

Jolsvai, H., McCauley, S.M., & Christiansen, M.H. (2013). Meaning Overrides Frequency in Idiomatic and Compositional Multiword Chunks. *Cognitive Science, 35*.

Keatley, C. W., Spinks, J. A., & de Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition*, 22, 70-84.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, 42(3):627–633.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language*, 49, 484–499.

Langacker, R.W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A Web-based interface for bilingual research. *Behavior Research Methods*, 38(2), 202–210.

Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.

Meyer, D., & Schvaneveldt, R. (1971). 'Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations,' *Journal of Experimental Psychology,* 90: 227–35.

Milton, J. (2010). The development of vocabulary breadth across the CEFR levels: A common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, and textbooks across Europe. In I. Bartning, M. Martin, & I. Vedder (Eds.), Communicative proficiency and linguistic development: Intersections between SLA and language testing research (pp. 211–232). Eurosla Monograph Series I.

Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, 56(1), pp. 1-24.

Nation, I.S.P. (2006). How Large a Vocabulary Is Needed for Reading and Listening? *The Canadian Modern Language Review / La revue canadienne des langues vivantes* 63(1), 59-81.

Nation, I.S.P. & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.

Öksüz, D. C. (2019). *Collocational processing in typologically different languages, English and Turkish: Evidence from corpora and psycholinguistic experimentation* (Doctoral dissertation). Retrieved from https://eprints.lancs.ac.uk/id/eprint/139672, Lancaster University.

Öksüz, D. C., Brezina, V., & Rebuschat, P. (2020). Collocational Processing in L1 and L2: The Effects of Word Frequency, Collocational Frequency, and Association. *Language Learning*. https://doi.org/10.1111/lang.12427

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology*, 55A, 1339- 1362.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114, 510–532.

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Press.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in Bilingualism and Second Language Learning. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (p. 371–388). Oxford University Press.

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W.J.B. (2011). Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Siyanova-Chanturia, A. (2013). Eye-tracking and ERPs in multi-word expression research: A state-of-the-art review of the method and findings. *Mental Lexicon*, 8(2), 245–268. http://doi.org/10.1075/ml.8.2.06siy

Stæhr, L.S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), 139–152.

Stubbs, M. (1995). Collocations and semantic profiles: on the cause of the trouble with quantitative methods. *Functions of language*, 2(1), 1-33.

The jamovi project. (2019). jamovi. (Version 0.9) [Computer Software]. Retrieved from https://www.jamovi.org.

van Hell, J.G., & Tanner, D. (2012). Second Language Proficiency and Cross-Language Lexical Activation. *Language Learning*, *62*(SUPPL. 2), 148–171.

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39(3), 458–483.

Winter, B. (2019). *Statistics for Linguists: An Introduction Using R* (1st ed.). Routledge. https://doi.org/10.4324/9781315165547

Wolter, B., & Gyllstad, H. (2011). Collocational Links in the L2 Mental Lexicon and the Influence of L1 Intralexical Knowledge. *Applied Linguistics*, 32(4), 430–449.

Wolter, B., & Gyllstad, H. (2013). Frequency of Input and L2 Collocational Processing. *Studies in Second Language Acquisition*, 35, 451–482.

Wolter, B., & Yamashita, J. (2014). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, 1(29), 1–29.

Wolter, B., & Yamashita, J. (2017). Word Frequency, Collocational Frequency, L1 Congruency, and Proficiency in L2 Collocational Processing. *Studies in Second Language Acquisition*, 1–22.

Yamashita, J., & Jiang, N. (2010). L1 Influence on the Acquisition of L2 Collocations: Japanese ESL Users and EFL Learners Acquiring English Collocations. *TESOL Quarterly*, *44*(4), 647–668.

Yamashita, J. (2018). Possibility of semantic involvement in the L1-L2 congruency effect in the processing of L2 collocations. *Journal of Second Language Studies*, *1*, 60–78.

Zhao, X., Li, P., Liu, Y., Fang, X., & Shu, H. (2011). Cross-language priming in Chinese-English bilinguals with different second language proficiency levels. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 801–806.

Zinkgraf, M. (2008). V+N Miscollocations in the Written Production of University Level Students, *ELIA*, 8, 91-116.