Ewa Dąbrowska and Tan Arda Gedik

Lexical knowledge, memory and experience

https://doi.org/10.1515/gcla-2023-0004

Abstract: Knowing a language involves knowing a large number of idiosyncratic units such as individual words and collocations which must be learned from the input. This study explores the role of implicit and explicit memory, as well as language experience, and specifically print exposure, with regard to adult native English speakers' knowledge of vocabulary and collocations. Consistent with prior research, our findings reveal a strong correlation between print exposure and performance on both language tasks. However, contrary to predictions, there were no significant effects of either implicit or explicit memory on either task. We argue that this is most likely due to the fact that language learning relies primarily on memory for associations between form and meaning and between words rather than memory sequences of meaningless phonological forms.

Keywords: collocations; vocabulary; individual differences; implicit memory; explicit memory; print exposure

1 Introduction

Learning a language involves learning many different types of units, including the forms and meanings of a large number of individual words and lexically specific phrases such as collocations. All this information is idiosyncratic: there is nothing about the phonological form/kæt/that would enable a person who does not know English or a related language to infer that it refers to a small furry purring animal with a tail. Similarly, there is no general principle that explains why it is idiomatic to say *strong wind* and *severe storm* but not *severe wind* or *weak wind* (cf. Herbst 2014). All this idiosyncrasy needs to be memorized.

This raises the question of how the learning actually happens and specifically, what memory systems are involved. Memory researchers distinguish between two distinct memory systems: explicit or declarative memory and implicit or procedural memory. Strictly speaking, the terms *explicit* and *declarative* on the one hand and

* **Corresponding author: Ewa Dąbrowska and Tan Arda Gedik**

implicit and *procedural* on the other are not strictly synonymous: the distinction between implicit and explicit memory is based on whether learners are consciously aware of what they learn, while the *declarative*/*procedural* distinction is based on the brain system involved. However, for the purposes of the present discussion we will use them interchangeably. Learning in explicit memory is conscious. It is also typically intentional, relatively fast, and depends on the hippocampus and other medial temporal lobe structures. Implicit learning, by contrast, occurs without conscious awareness, is effortless and incidental (that is to say, it occurs without the intention to learn) and relatively slow (i.e. it requires a substantial amount of repetition). The neural circuits involved include the basal ganglia and frontal cortex (see Nicolson et al. 2010 and Ullman 2004, 2016 for further discussion).

There is no consensus on what role these memory systems play in language learning. It is almost universally assumed that learning words and their meanings depends on the explicit memory system (Lum and Conti-Ramsden 2013; Hamrick et al. 2018). This is because the resulting knowledge is arguably explicit: for example, speakers are aware that the phonological form *cat* refers to a particular kind of animal. Furthermore, adults and older children are able to learn new vocabulary items from verbal definitions, and learning the form-meaning pairing of a word has been observed to occur relatively fast, sometimes after a single exposure (Carey and Bartlett 1978; Golinkoff et al. 1992). Last but not least, a number of studies have shown that vocabulary size correlates with explicit memory (see Hamrick et al. 2018 for a recent review). In fact, some common explicit memory tasks (for example, when participants have to learn associations between a picture and a phonological form) could be regarded as vocabulary learning tasks.

On the other hand, there is some evidence that implicit memory may also be involved in vocabulary learning. Research on cross-situational learning has shown that learners are able to track co-occurrence statistics between words and referents across multiple ambiguous situations (see e.g. Smith and Yu 2008; Smith et al. 2011) and it has been proposed that this is the mechanism underlying word learning in infants (Pereira et al. 2013; Vlach and Johnson 2013). Since cross-situational learning has been observed in amnesic patients, who have severely impaired declarative memory, it is generally assumed that the procedural system is involved in such contexts. However, other studies suggest that explicit hypothesis testing plays a role here, too (Kachergis et al. 2010). It has also been argued that lexical knowledge may be implicit in some cases. Dąbrowska (2014a) found that subjects perform well above chance levels when they are given a multiple-choice task in which they are asked to guess the meanings of words they claim they do not know. According to Dienes (2008), knowledge can be considered implicit when participants believe they are guessing but their performance is actually above chance. On this criterion, therefore, Dąbrowska's

(2014a) participants showed implicit knowledge of word meanings which was presumably acquired using implicit memory.

Little is known about which memory system is involved in learning and storing collocations. There is some evidence that collocation learning relies on implicit learning mechanisms. Firstly, naïve native speakers are not aware of collocations (at least until they hear collocational norms violated; cf. Eyckmans 2009). Secondly, to the extent that collocations can be defined in statistical terms, learning them involves tallying of statistical co-occurrence patterns of the words that make them up (Forsberg Lundell and Sandgren 2013). Finally, Yi (2018) found that implicit memory (as measured by a serial reaction time task) predicted the processing speed for multi-word sequences in native speakers.

On the other hand, Ullman (2016) argues that collocations should depend on explicit memory as they are part of the lexicon and the lexicon is learned explicitly. There is some empirical evidence to support this: Llompart and Dąbrowska (2020) found a significant relationship between collocations and explicit memory (assessed by a paired associates task in which participants had to learn the association between novel shapes and nonce words) but not between collocations and implicit memory (assessed using a serial reaction time task). However, since the explicit memory task utilized verbal stimuli while the SRT task was entirely non-verbal, this finding could be an artifact of the testing methods.

Thus, earlier research on this topic is inconclusive. In this study, therefore, we examine the possible effects of both implicit and explicit memory on adult native speakers' knowledge of English vocabulary and collocations, using verbal stimuli in both memory tasks. We also administer a measure of print exposure in order to measure the quality of linguistic experience. Since knowledge of both vocabulary and collocations is known to correlate with print exposure (Brown 2021; Dąbrowska 2018; Mol and Bus 2011; Stanovich and Cunningham 1992), including a measure of the latter will allow us to estimate the extent to which individual differences in these two areas of linguistic knowledge depend on experience as opposed to learner-internal characteristics such as implicit and explicit memory.

2 Method

2.1 Participants

Thirty-eight L1 English speaking participants (25 females) took part in the study in exchange for financial compensation. Participants' mean age was 41.1 (SD 8.8, range 23–60) and the mean number of years spent in formal education was 16.0

(SD 3.2, range 11–23). They were recruited via Prolific, an online participant recruitment platform (www.prolific.com) and tested online. Participants gave their consent to participate in the study in accordance with the Declaration of Helsinki.

2.2 Materials

The study consisted of two cognitive tasks, two language tasks and a measure of print exposure. Tasks are described in the same order as in the study. Participants also completed a questionnaire on their background, in which they were asked questions about their L1, any other languages they spoke, age, occupation, number of years in formal education, highest academic attainment and how much they read on average.

2.2.1 Implicit Memory for Syllable Sequences

Implicit learning tasks are notoriously unreliable. As argued by Siegelman et al. (2017), this is due to several reasons: the number of trials in the test phase is often too small, the test items are typically of the same level of difficulty, and a high proportion of participants perform at chance level. Because of this, most existing implicit learning tasks are suitable for group-level comparisons but not sensitive enough to reliably detect individual differences in implicit learning ability. Siegelman et al. (2017) therefore developed a new visual implicit statistical learning task which was specifically designed for use in studies investigating individual differences. The task is based on the "embedded triplets" task used by Frost et al. (2013) but contains more test items which differ in difficulty. This was achieved by systematically manipulating transitional probabilities within triplets, sequence length and the number of distractors in the test phase. As expected, the task has much better psychometric properties: a Cronbach's alpha of .88, a split-half reliability of .83, and a test–retest reliability of .68. Our task was closely modelled on theirs but contained sequences of syllables instead of visual stimuli.

The task consisted of a familiarization phase and a test phase. During the familiarization phase, participants were presented with a continuous sequence of 16 syllables organized into 8 triplets. Four triplets were constructed from four syllables arranged in four different orders (*righotu, tusegho, seturi, ghorise*) and had transitional probabilities between syllables of .33. The other four triplets (*bovnuskwa, fuwoyi, skwuziyo, smoyusi*) were constructed by concatenating the remaining 12 syllables into four sequences of three; the transitional probabilities in these triplets were therefore 1. The order of the triplets was semi-random, with the constraint that the same triplet could not be repeated twice in a row. The familiarization phase consisted of 24 blocks, where a block consisted of a single presentation of all eight triplets. Following Kidd and Arciuli (2016), we added a cover task intended to check if the participants were paying attention during the familiarization phase. At the end of each block, an additional triplet was inserted in which one of the syllables was repeated twice (e.g., *righotutu*). Participants were instructed to press the spacebar as quickly as possible when they heard the same syllable repeated.

The test phase consisted of 42 forced-choice trials. In each trial participants heard either a triplet or a part triplet (the first two or the last two syllables of a triplet) and either one or three distractors. The distractors were created by concatenating sequences of syllables belonging to different triplets (see Siegelman et al. 2017 for details). The participants' task was to choose the sequence that they heard during the familiarization phase. At the beginning of each trial, the target and the distractors played one by one. As each sequence played, a number appeared at a different location on the screen (1 in the upper left hand quadrant, then 2 in the upper right hand quadrant, and so on). When all the sequences had played, all numbers appeared on the screen at the same time and the participants were asked to click on the number corresponding to the sequence that sounded familiar (see Figure 1). Participants could listen to each stimulus again by clicking on the corresponding number as many times as they needed to before responding. The position of the target was counterbalanced across test items.

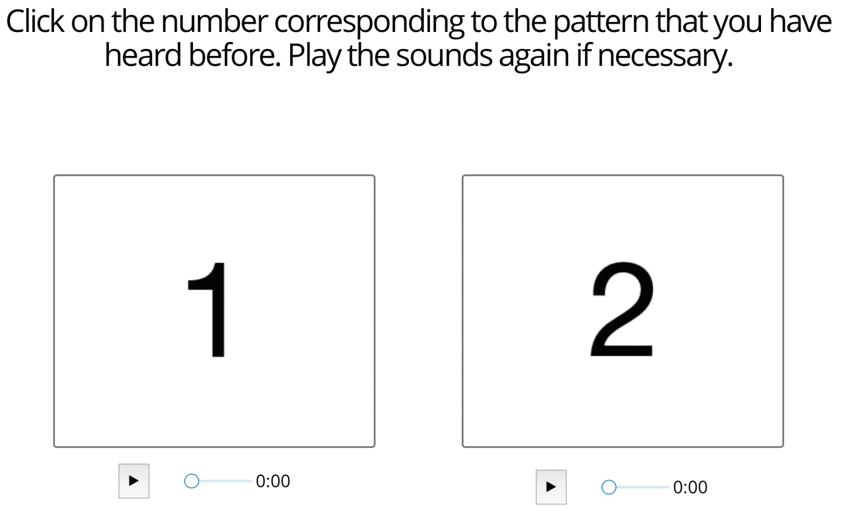

Figure 1: Screenshot of the test phase of the implicit memory task.

In addition to using syllables instead of visual stimuli, our version of the task differed from that used by Siegelman et al. (2017) in three ways. Firstly, the original version of the task had nine test items with fill-in-the-blank questions. However, piloting revealed that this format is extremely difficult in the auditory modality, so these items were replaced with standard forced-choice items. Secondly, as explained above, we added a cover task during the familiarization phase in order to ensure that the participants remained attentive. Finally, at the end of the familiarization phase, participants were asked whether they noticed any repeated sequences, and, if so, how long the sequences were (2, 3, 4 or 6 syllables). This was done in order to check whether participants were consciously aware of the repeated sequences.

The syllable stimuli were created on voicemaker.in (www.voicemaker.in) using a general US male accent at 24000hz sample rate. Voicemaker is a free website designed to generate natural sounding language output in a variety of languages. The average duration of each syllable was 500ms and a new syllable was presented every 800ms.

2.2.2 Collocations

Participants' knowledge of collocations was measured using an expanded version of the Words that Go Together Test (Dąbrowska 2014b). The original test contains 40 multiple-choice questions, each consisting of five short phrases. One of the phrases is an established collocation; the other four are plausible distractors. The distractor items were semantically close to the target item, but their mutual information scores were much lower (e.g. *blatant lie – clear lie – conspicuous lie – distinct lie – recognizable lie*; the target item in this case is *blatant lie*). The participants were asked to choose the phrase that "sounds the most natural or familiar". In addition to the original 40 items, 9 extra items were sourced from Garibyan (2022). Garibyan's task was closely modelled on Dąbrowska's but consisted of 60 items which were somewhat easier, as they were intended for advanced second language learners rather than native speakers. Since Dąbrowska's original test is quite difficult, in the extended version used in the current study we incorporated the nine additional items.[1]

1 The new items were *boost morale, joint effort, nagging doubt, slim chance, abuse trust, feeble excuse, suffer casualties, casual acquaintances and hurl insults*.

2.2.3 Author Recognition Test

To measure print exposure, we used the author recognition test (ART; see Acheson et al. 2008). The test consists of a list of 130 names, including 65 authors' names and 65 foils. The participants' task is to select those names which belong to real authors. Previous research has shown that the ART is a reliable measure of print exposure and usually provides a more accurate estimate than self-report questionnaires, in which participants tend to give socially desirable answers. In our computerized version of the task, names were presented on the screen one by one, and participants were asked to click either 'AUTHOR' or 'DON'T KNOW'. The score was the number of times participants correctly identified real authors and rejected foil items (maximum possible score = 130).

2.2.4 Vocabulary

An abridged version of the vocabulary size test (Nation and Beglar 2007) was computerized and implemented in the study to measure English vocabulary knowledge. The abridged version was adapted from Dąbrowska (2018) and consists of 60 multiple-choice items systematically sampled from various frequency bands. Vocabulary items were presented on screen one at a time. In each trial, the target item was presented in uppercase letters at the top of the screen. Four short definitions were printed below. Participants were instructed to select the most appropriate definition. The score was the total number of correct choices. The maximum possible score is thus 60.

2.2.5 Explicit Memory for Syllable Sequences

This task was similar to the implicit memory task described above, but the instructions given to the participants were different. Participants were told that they would be presented with sequences of three syllables, and that the task was to memorize these sequences. 24 new syllables (different from the ones used in the implicit memory task) were created on voicemaker.in using the same settings as before and combined into 8 triplets. During the familiarization phase, each triplet was followed by a 250 ms pause. A fixation cross was displayed on the screen during the pause. Since learning under such conditions is much faster, there were only 6 training blocks. The cover task and the testing phase were exactly the same as in the implicit memory task.

2.3 Procedure

All tasks were created, piloted, and implemented on Gorilla, an online experimental platform (www.gorilla.sc), and are freely available to all researchers on Gorilla open materials (https://app.gorilla.sc/openmaterials/667440).

Participants were first instructed to read the consent form and give their consent. Then, they were presented with the background questionnaire. Following this was a sound check task to make sure their computers had working speakers or headphones, and their browsers could play sounds automatically. The tasks were administered in the order described above. The entire experiment took approximately 38 minutes.

3 Results

The data were imported to RStudio (R Core Team 2023) after pre-processing to extract individual values for statistical analyses. Accuracy on all tasks was calculated by adding up the number of correct answers. The lmg in relaimpo package (see Grömping 2007) was used to estimate the relative importance of each predictor.

Table 1 provides information about performance on individual tasks as well as their reliabilities, while the distributions of individual scores are shown in Figure 2. As we can see, individual performance differs considerably on all tasks, and all tasks have reasonable reliabilities.

Table 1: Performance on individual tasks (raw scores).

Task	Mean (SD)	Range	Split-Half Reliability
Vocabulary	49.37 (5.60)	33–56	.78
Collocations	37.50 (6.17)	22–48	.84
ART	80.94 (9.42)	64–102	.92
Implicit memory	21.32 (3.66)	12–28	.77
Explicit memory	26.16 (3.81)	19–38	.75

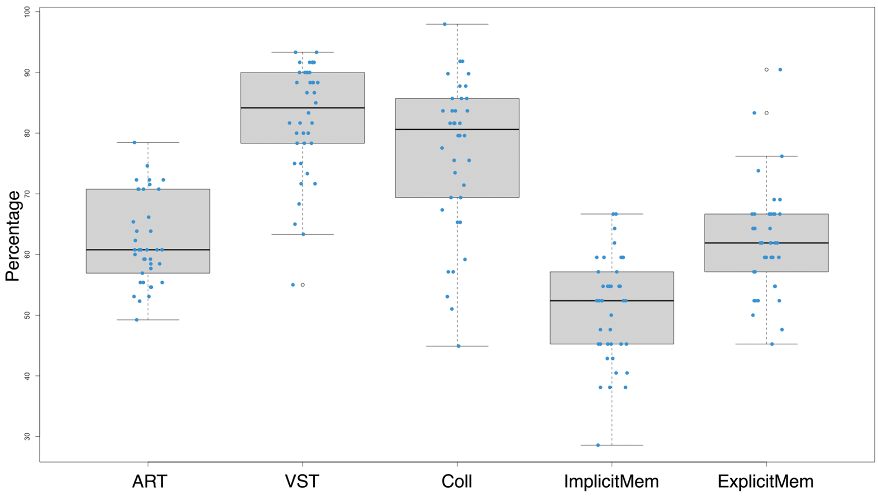

Figure 2: Performance on individual tasks (percent correct). Abbreviations: ART (author recognition test); VST (vocabulary size test); Coll (collocations test); ImplicitMem (implicit memory for syllables task); ExplicitMem (explicit memory for syllables task).

Table 2 summarizes participants' responses to the awareness question at the end of the familiarization phase in the implicit memory task. As we can see, 12 participants said that they did not notice a sequence. The remaining responses were distributed fairly randomly between the other 4 response options, suggesting that participants were simply guessing. This confirms that the implicit memory task is in fact implicit.

Table 2: Number of times participants selected what sequence they were presented with in the implicit memory task.

Sequence Type	Number
No fixed sequences of syllables	12
Some fixed sequences of 2 syllables	5
Some fixed sequences of 3 syllables	5
Some fixed sequences of 4 syllables	6
Some fixed sequences of 6 syllables	10

Table 3 provides information about correlations between all the variables, especially the language tasks and the memory tasks. As we can see, there are robust correlations between ART, vocabulary and collocations. None of the other correlations are significant.

Table 3: Correlation matrix for the two language tasks (Vocabulary and Collocations), the two memory tasks (Implicit Memory and Explicit Memory), the print exposure measure (ART) and the number of years spent in formal education (Education).

	Vocabulary	Collocations	ART	Implicit Memory	Explicit Memory	Education
Vocabulary	1.00					
Collocations	.58***	1.00				
ART	.67***	.44**	1.00			
Implicit memory	.10	.18	.01	1.00		
Explicit memory	.17	.14	.11	.27	1.00	
Education	.24	.20	.16	−.09	.09	1.00

**significant at* $p < .05$, ***significant at* $p < .01$, ****significant at* $p < .001$.

Tables 4 and 5 present summaries of ordinary least square regression models predicting performance on the Vocabulary and Collocations from the ART, implicit memory and explicit memory tasks. The models confirm what was already obvious from the correlational analysis, namely that ART is the only significant predictor of performance on the language tasks, accounting for over 44% of the variance in scores on the vocabulary task and just under 19% of the variance in the collocations task.

Table 4: Regression results for Vocabulary.

Variable	Estimate	Std. Error	*t* value	Pr(>|t|)	lmg
intercept	11.92579	7.93362	1.503	0.142	
ART	0.39508	0.07477	5.284	.000***	0.4443
Implicit Memory	0.12131	0.19822	0.612	0.545	0.0078
Explicit Memory	0.11155	0.19171	0.582	0.564	0.0166
Model R^2					0.4689

Table 5: Regression results for Collocations.

Variable	Estimate	Std. Error	*t* value	Pr(>|t|)	lmg
intercept	6.67535	10.55460	0.632	0.53132	
ART	.028407	0.09947	2.856	0.00727**	0.1885
Implicit Memory	0.28220	0.26370	1.070	0.29208	0.0289
Explicit Memory	0.07050	0.25505	0.276	0.78390	0.009
Model R^2					0.2269

****significant at p < .01.*

4 Discussion

We found robust relationships between performance on the two language tasks, as well as performance on the language tasks and print exposure. The finding that vocabulary correlates with print exposure is well established (see, for example, Cunningham and Stanovich 1998; Dąbrowska 2018; James et al. 2018). The relationship between collocations and print exposure has also been observed in earlier research (see, for example, Dąbrowska 2018; Llompart and Dąbrowska 2020). The strong relationship between vocabulary and collocations is compatible with both the declarative procedural model (Ullman 2016) and Dąbrowska's (2009) suggestion that tracking a word's collocations and semantic preferences is an important source of information about its meaning.

On the other hand, there was no significant relationship between performance on the language tasks and either of the memory tasks. The two memory tasks are also not correlated with each other, suggesting that, as intended, they measure different abilities, thus providing divergent construct validity for both tasks.

The non-effect of either implicit or explicit memory is unexpected given the considerations outlined earlier in this paper: both the relationship between form and meaning and the relationship between collocates are arbitrary, that is to say, not predictable from general principles, and therefore need to be stored in memory. Given that there are considerable individual differences in performance on both language tasks on the one hand, and both memory tasks on the other, and that all the tasks investigated in this study have reasonable reliabilities, we should have found a robust correlation between the language measures and at least one of the memory measures.

Regarding vocabulary, it could be argued that any existing relationship with memory could be masked by a much stronger relationship with other, stronger, predictors, in particular print exposure and IQ. The strong correlation with print exposure is due to the fact that printed text contains a much higher proportion of infrequent words than spoken language (cf. Cunningham and Stanovich 1998); therefore, once learners have acquired a basic vocabulary, they simply do not encounter many unfamiliar words in spoken language. The strong correlation between vocabulary size and IQ (cf. Bloom 2000; Jensen 2001; Sternberg 1987) is due to the fact that in order to learn words, one has to infer their meanings from the situational and/or linguistic context (cf. Sternberg 1987), and arguably the mental processes required to do this are the same as those measured by IQ tests.

However, no such explanation can be offered for collocations. Learning which words co-occur does not require inferencing but simply storing information about the contingencies of language use. The explanation we would like to propose for the lack of a relationship between collocations and either of our memory tasks is that the latter measured memory for phonological forms, but what is more relevant for language learning is memory for associations between words and between form and meaning. After all, both words and collocations are symbolic forms, that is to say, form-meaning pairings.

Our argument is based on three well-known findings from psychology. Firstly, it is well established that lexical access is obligatory. That is to say, when a speaker hears a phonological form such as *cat*, they automatically access its meaning. Evidence for this comes from a host of priming studies, including studies demonstrating obligatory access of contextually irrelevant meanings (e.g. Swinney 1979). Likewise, when speakers hear a complex expression, they obligatorily access its meaning: for example, the phonological form *black cat* automatically evokes the concept of a cat with black fur.

Secondly, memory for meaningful linguistic expressions is much more accurate than memory for form alone. For instance, speakers remember the meaning of a word better than its phonological form or the physical characteristics of the printed form. This is true for both recognition and recall, and for intentional as well as incidental learning (Craik and Tulving 1975). Similarly, speakers remember the gist (i.e. the main point) of a sentence much better than the form (Sachs 1967). Furthermore, lists of words are remembered better than lists of nonwords (i.e. word-like phonological forms), even if participants have been familiarized with these forms (Hulme et al. 1995).

The final relevant piece of evidence is the fact that the use of mnemonics is an effective strategy in learning the phonological forms of foreign language words (Hulstijn 1997; Wei 2015). The most relevant mnemonic technique is the use of linkwords – words that have the same, or similar, pronunciation in a known language as the target word – and forming conceptual associations between the meaning of the linkword and the meaning of the target word. For example, in order to learn the German word for ‘table’, *Tisch*, one could think of the English word *dish* and imagine a sandwich served on a miniature table; or in order to learn the Spanish word for ‘to be’, *estar*, one could think of the phrase *you are a star*. Note that when using the linkword technique, a learner must remember more than the simple association between two phonological forms (*table* = *Tisch*; *to be* = *estar*). The technique is effective because it turns learning a meaningless association into learning a meaningful one.

5 Conclusion

This study investigated the effects of implicit and explicit verbal memory as well as print exposure on vocabulary and collocational knowledge among adult native speakers of English. In line with previous research, we found a robust effect of print exposure on both language measures. However, there were no significant effects of either implicit or explicit memory for syllable sequences. We have argued that this is because what matters for language learning is memory for co-occurrences of symbolic units (that is to say, form-meaning pairings) rather than phonological forms.

This conclusion may not come as a surprise to researchers working in the cognitive linguistics/construction grammar tradition, who see language as an assembly of symbolic units, i.e. form-meaning pairings (Goldberg 2006, Langacker 1987). However, many other traditions in linguistics, including not just formal approaches, but also work in the artificial grammar paradigm and most of the research investigating the role of implicit learning in language acquisition assume that learning a language involves, in essence, learning about relationships between units defined at the level of form (cf. Chomsky 1965, Petersson et al. 2012, Reber 1967). It is possible, of course, that such relationships are more relevant at earlier stages of acquisition. Nevertheless, our results suggest that future research cannot simply assume that learners attend just to co-occurrence relationships between forms, and we must seriously consider the possibility that the semantic relationships between the meanings that these forms convey are more relevant for acquisition.

Understanding what types of mental processes are involved in building a mental constructicon, and the types of units they operate on (e.g. phonological forms versus form-meaning pairings as well as their size and degree of specificity) is crucial for theories of language which adopt the cognitive commitment (cf. Lakoff 1991: 54). However, many cognitive linguists merely gesture towards the latter and fail to engage with the relevant psycho- and neurolinguistic questions (cf. Dąbrowska 2016). We hope that this modest contribution to research on the mental processes involved in acquiring lexical constructions will help to convince our fellow linguists that these are important issues, and that some assumptions that appear obvious may not be correct.

References

Acheson, Daniel J., Justine B. Wells & Maryellen C. MacDonald. 2008. New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods* 40(1). 278–289. DOI: 10.3758/BRM.40.1.278.

Bloom, Paul. 2000. *How children learn the meanings of words*. Cambridge, MA: The MIT Press.

Brown, Jessica. 2021. *Exploring the relationship between socioeconomic status, language exposure, and language processing in young adults*. York: York St John University Doctoral Dissertation. https://ray.yorksj.ac.uk/id/eprint/5386.

Carey, Susan & Elsa Bartlett. 1978. Acquiring a single new word. *Papers and Reports on Child Language Development* 15. 17–29.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.

Craik, Fergus I. M. & Endel Tulving. 1975. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104(3). 268–294. DOI: 10.1037/0096-3445.104.3.268.

Cunningham, Anne E. & Keith E. Stanovich. 1998. What reading does for the mind. *American Educator* 22. 8–17.

Dąbrowska, Ewa. 2009. Words as constructions. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New Directions in Cognitive Linguistics*, 201–223. Amsterdam: John Benjamins.

Dąbrowska, Ewa. 2014a. Implicit lexical knowledge. *Linguistics* 52(1). 205–223. DOI: 10.1515/ling-2013-0060.

Dąbrowska, Ewa. 2014b. Words that go together: Measuring individual differences in native speakers' knowledge of collocations. *The Mental Lexicon* 9(3). 401–418. DOI: 10.1075/ml.9.3.02dab.

Dąbrowska, Ewa. 2016. Cognitive linguistics' seven deadly sins. *Cognitive Linguistics* 27(4). 479–491. DOI: 10.1515/cog-2016-0059.

Dąbrowska, Ewa. 2018. Experience, aptitude and individual differences in native language ultimate attainment. *Cognition* 178. 222–235. DOI: 10.1016/j.cognition.2018.05.018.

Dienes, Zoltán. 2007. Subjective measures of unconscious knowledge. *Progress in Brain Research* 168. 49–64.

Eyckmans, June. 2009. Toward an assessment of learners' receptive and productive syntagmatic knowledge. In Andy Barfield & Henrik Gyllstad *(eds.), Researching Collocations in Another Language*, 139–152. London: Palgrave Macmillan. DOI: 10.1057/9780230245327_11.

Forsberg Lundell, Fanny & Maria Sandgren. 2013. High-level proficiency in late L2 acquisition: Relationships between collocational production, language aptitude and personality. In Gisela Granena & Mike Long (eds.), *Sensitive periods, language aptitude, and ultimate*

L2 attainment (Language Learning & Language Teaching, 35), 231–256. Amsterdam & Philadelphia: John Benjamins. DOI: 10.1075/lllt.35.09for

Frost, Ram, Noam Siegelman, Alona Narkiss & Liron Afek. 2013. What predicts successful literacy acquisition in a second language? *Psychological Science* 24(7). 1243–1252. DOI: 10.1177/0956797612472207.

Garibyan, Armine. 2022. *Processing collocations cross-linguistically: A psycholinguistic and neurolinguistic study*. Erlangen & Nuremberg: Friedrich-Alexander-Universität Erlangen-Nürnberg Doctoral Dissertation. urn:nbn:de:bvb:29-opus4-211773

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Golinkoff, Roberta M., Kathy Hirsh-Pasek, Leslie M. Bailey & Neill R. Wenger. 1992. Young children and adults use lexical principles to learn new nouns. *Developmental Psychology* 28(1). 99–108. DOI: 10.1037/0012-1649.28.1.99.

Grömping, Ulrike. 2007. Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician* 61(2). 139–147. DOI: 10.1198/000313007X188252.

Hamrick, Phillip, Jarrad A. G. Lum & Michael T. Ullman. 2018. Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences* 115(7). 1487–1492. DOI: 10.1073/pnas.1713975115.

Herbst, Thomas. 2014. Idiosyncrasies and generalizations: Argument structure, semantic roles and the valency realization principle. *Yearbook of the German Cognitive Linguistics Association* 2(1). 253–289. DOI: 10.1515/gcla-2014-0015.

Hulme, Charles, Steven Roodenrys, Gordon Brown & Robin Mercer. 1995. The role of long-term memory mechanisms in memory span. *British Journal of Psychology* 86(4). 527–536. DOI: 10.1111/j.2044-8295.1995.tb02570.x.

Hulstijn, Jan H. 1997. Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In James Coady & Thomas Huckin (eds.), *Second language vocabulary acquisition: A rationale for pedagogy*, 203–224. Cambridge: Cambridge University Press.

James, Ariel N., Scott H. Fraundorf, Eun-Kyung Lee & Duane G. Watson. 2018. Individual differences in syntactic processing: Is there evidence for reader-text interactions? *Journal of Memory and Language* 102. 155–181. DOI: 10.1016/j.jml.2018.05.

Jensen, Arthur R. 2001. Vocabulary and general intelligence. *Behavioral and Brain Sciences* 24(6). 1109–1110. DOI: 10.1017/S0140525X01280133.

Kachergis, George, Chen Yu & Richard Shiffrin. 2010. Cross-situational statistical learning: Implicit or intentional? In *Proceedings of the Annual Meeting of the Cognitive Science Society* 32. 2362–2367.

Kidd, Evan & Joanne Arciuli. 2016. Individual differences in statistical learning predict children's comprehension of syntax. *Child Development* 87(1). 184–193. DOI: 10.1111/cdev.12461.

Lakoff, George. 1991. Cognitive vs generative linguistics: How commitments influence results. *Language & Communication* 11(1–2). 53–62. DOI: 10.1016/0271-5309(91)90018-Q.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar*. Vol. 1, *Theoretical Prerequisites*. Stanford: Stanford University Press.

Llompart, Miguel & Ewa Dąbrowska. 2020. Explicit but not implicit memory predicts ultimate attainment in the native language. *Frontiers in Psychology* 11. DOI: 10.3389/fpsyg.2020.569586.

Lum, Jarrad A. G. & Gina Conti-Ramsden. 2013. Long-term memory: A review and meta-analysis of studies of declarative and procedural memory in specific language impairment. *Topics in Language Disorders* 33(4). 282–297. DOI: 10.1097/01.TLD.0000437939.01237.6a.
Mol, Suzanne E. & Adriana G. Bus. 2011. To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin* 137(2). 267–296. DOI: 10.1037/a0021890.
Nation, Paul & David Beglar. 2007. A vocabulary size test. *The Language Teacher* 31(7). 9–13.
Nicolson, Roderick I., Angela J. Fawcett, Rebecca. L. Brookes & Jamie Needle. 2010. Procedural learning and dyslexia. *Dyslexia* 16(3). 194–212. DOI: 10.1002/dys.408.
Pereira, Alfredo F., Linda B. Smith & Chen Yu. 2013. A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review* 21(1). 178–185.
Petersson, Karl-Magnus, Vasiliki Folia & Peter Hagoort. 2012. What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language* 120(2). 83–95. DOI: 10.1016/j.bandl.2010.08.003.
R Core Team. 2023. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.r-project.org/.
Reber, Arthur S. 1967. Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior* 6(6). 855–863. DOI: 10.1016/S0022-5371(67)80149-X.
Sachs, Jacqueline Strunk. 1967. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics* 2(9). 437–442. DOI: 10.3758/BF03208784.
Siegelman, Noam, Louisa Bogaerts & Ram Frost. 2017. Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods* 49(2). 418–432. DOI: 10.3758/s13428-016-0719-z.
Smith, Kenny, Andrew D. M. Smith & Richard A. Blythe. 2011. Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science* 35(3). 480–498. DOI: 10.1111/j.1551-6709.2010.01158.x.
Smith, Linda & Chen Yu. 2008. Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition* 106(3). 1558–1568. DOI: 10.1016/j.cognition.2007.06.010.
Stanovich, Keith E. & Anne E. Cunningham. 1992. Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition* 20(1). 51–68. DOI: 10.3758/BF03208254.
Sternberg, Robert J. 1987. Most vocabulary is learned from context. In Margaret G. McKeown & Mary E. Curtis (eds.), *The nature of vocabulary acquisition*, 89–106. Hillsdale: Lawrence Erlbaum Associates.
Swinney, David A. 1979. Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 18(6). 645–659. DOI: 10.1016/S0022-5371(79)90355-4.
Ullman, Michael T. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92(1–2). 231–270. DOI: 10.1016/j.cognition.2003.10.008.
Ullman, Michael T. 2016. The declarative/procedural model. In Gregory Hickok & Steven L. Small (eds.), *Neurobiology of language*, 953–968. Amsterdam: Elsevier. DOI: 10.1016/B978-0-12-407794-2.00076-6.
Vlach, Haley A. & Scott P. Johnson. 2013. Memory constraints on infants' cross-situational statistical learning. *Cognition* 127(3). 375–382. DOI: 10.1016/j.cognition.2013.02.015.

Wei, Zheng. 2015. Does teaching mnemonics for vocabulary learning make a difference? Putting the keyword method and the word part technique to the test. *Language Teaching Research* 19(1). 43–69. DOI: 10.1177/1362168814541734.

Yi, Wei. 2018. Statistical sensitivity, cognitive aptitudes, and processing of collocations. *Studies in Second Language Acquisition* 40(4). 831–856. DOI: 10.1017/S0272263118000141.