UDK 81’37
81’232
Izvorni znanstveni rad
Rukopis primljen 21. IX. 2024.
Prihvaćen za tisak 12. VI. 2025.
https://doi.org/10.31724/rihjj.51.1.7

Anita Peti-Stantić
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, HR-10000 Zagreb
https://orcid.org/0000-0001-5650-5719
anita.peti-stantic@ffzg.hr

SOCIAL CONSTRUCTION OF AN ABSTRACT LEXICON

Abstract word meanings, being purely mental constructs, depend on the mediation of language to be learnt, as well as recalled and used. Recent findings showed that adult native speakers of the same language do not share the same mental grammar and vocabularies because they attend to different cues in the input. Therefore, we conducted a study investigating the role of educational attainment and exposure to rich reading experiences (RRE) as social factors impacting the capacity to form associations and recall collocations within the abstract lexicon. The primary objectives of this study were: 1. to establish a relationship between social factors and the capacity to form associations and recall collocations within abstract lexicon and, 2. to understand the parallelism between social factors affecting the potential for building abstract lexicon and the types of cues encoded in the architecture of language. We not only found that two populations differ in the size and depth of their abstract mental lexicons, but also that there is a parallelism between the types of cues conditioned by the education attainment and the exposure to RRE.[1]

1. Introduction

Languages, understood as social tools, play an important societal role in integrating the body of knowledge dynamically distributed across individuals and populations (Borghi anf Binkofski, 2014). In doing so, they refer to the world available to our senses (mirrored in concepts with material referents, represent-

[1] The paper was written as part of the NPOO project The Social Construction of Abstract Meanings through Reading, PI Anita Peti-Stantić.

ed by concrete words), but also to the world available only through our cognitive systems (mirrored in concepts with no material referents, represented by abstract words). A crucial difference between them is that abstract word meanings, being purely mental constructs, depend on the mediation of language for their acquisition, recall and use. The distinction in processing of these two subsets has been well-documented. The concreteness effect was reported in a number of experimental tasks and paradigms (e.g., De Groot, 1989; De Groot anf Keijzer, 2000; Binder et al., 2005; Fliessbach et al., 2006; Romani et al., 2008), as well as in reaction time (RT) and an event-related potential (ERP) studies (e.g., Schwanenflugel et al., 1992; Kounios anf Holcomb, 1994; West anf Holcomb, 2000; Barber et al., 2013). However, the study of linguistic and extralinguistic cues relevant for the emergence and use of abstract words in distinct populations is missing.

In this paper we build on two assumptions: 1. language use in education and reading practice support vocabulary growth through probabilistic multiple-cue learning processes (Gleitman et al., 2005); 2. levels of education and exposure to rich reading experiences (RRE) (Cevoli et al., 2022; Castles et al., 2018) are responsible for large individual differences in vocabulary size, collocations, and grammars of adult speakers (Dąbrowska, 2012, 2014, 2015). We examine the capacity to process abstract words within the framework of social conditioning, particularly focusing on diverse educational backgrounds and exposure to RRE. Since it has been shown that the difference in conceptual content between verbs (referring to relations) and nouns (typically referring to objects), followed by grammatical valence, impacts their concreteness and therefore processing (for conceptual and neurocognitive differences see Langacker, 1987; Croft, 1991; Ma et al., 2009; Simonsen et al., 2013; Vigliocco et al., 2011; Lee anf Federmeier, 2008; for acquisition differences see Gentner, 1982; Nelson, Hampson anf Shaw, 1993; Tomasello et al., 1997), two minimal semantic pairs of complex prefixed verbs (CPV) were chosen for this study as an example of the most abstract words.[2]

[2] Since this study is the first of its kind, it should be considered preliminary. Reviewers noted that the study's methodological choices such as the number of participants, demographic homogeneity of the sample, as well as no control over socioeconomic status, parental education, cognitive abilities, and motivational factors may limit the results' generalizability. The study's results do nonetheless point at direction for further investigation.

Verbal pairs chosen for the study were *pripisati* 'to attribute' and *pridati*, 'to add, to attribute', and *procijeniti* 'to appraise' and *prosuditi* 'to judge'. These pairs were chosen because both root verbs *pisati* 'to write' and *dati* 'to give' are very frequent in Croatian, but different in meaning, as well as semantic scope and specificity. However, prefixed and formed as CPVs, they are similar both in meaning and in scope, yet different in collocations. Contrary to that, the root verbs *cijeniti* 'to appreciate' and *suditi* 'to judge' are far less frequent in Croatian, and also different in meaning. Prefixed, they are similar in meaning and in scope, but differ in collocations. Furthermore, emotional valence (very low) is controlled for in all four CPVs, as results could be dependent on high emotional valence (Kousta et al., 2011). We selected these pairs not to test a wide range of verbs, but rather to provide model examples that could direct future studies.

Furthermore, the analysis of the data in the Croatian Psycholinguistic Database (CPD) showed that the reported corpus frequency differs from subjective frequency more for abstract than for concrete words (Peti-Stantić et al., 2021). Therefore, we hypothesize that the subjective frequency of abstract lexicon, reflecting the density of the network, impacts the mechanisms of learning and recalling the meanings of abstract words.

The primary objectives of this study were: 1. to establish a relationship between social factors, such as educational attainment and exposure to RRE, and the capacity to form associations and recall collocations within abstract lexicon, and 2. to understand the parallelism between social factors affecting the potential for building abstract lexicon and the types of cues encoded in the architecture of language.

2. On concrete and abstract lexicons

Not only do many words not refer to anything perceptible, but so do sentences and constructions. Rather, these words, constructions and sentences are lexicalized concepts constructed by our minds (Jackendoff, 1994). Consequently, the category of concreteness provides an appropriate framework for examining patterns related to the distinctions between concrete and abstract words.

The referents of concrete words are usually perceptible or observable through the senses. The speakers encounter and utilize concrete words more often than the abstract ones because they are exposed to them from an early age. This gives the speakers more opportunities to learn them during their language experiences. Aspects like age of acquisition and prevalence of these words and constructions in everyday usage have an impact on both the subjective perception of their frequency and their objective frequency in the language corpus. Additionally, comprehending, remembering, and retaining the meanings of those words is easier because of the frequent interaction with the things they refer to. The primary frameworks that have been developed to help clarify the processing distinctions between concrete and abstract words are the context availability theory (Schwanenflugel anf Shoben, 1983), the dual-coding theory (Paivio, 1986, 2007), and grounding theories (e.g. Barsalou anf Wiemer-Hastings, 2005; Barsalou, 2008). In short, each of these three theories suggests a unique role for verbal/linguistic knowledge in producing concreteness effects, in addition to experiential knowledge. Concrete words are directly understood through our perceptual experience and physical interaction with them. This phenomenon is commonly referred to in the literature as the concreteness effect. This effect in psycholinguistics–where concrete words are processed more efficiently than abstract ones–has long been explained by dual-coding theory, which posits that concrete words benefit from both verbal and imagery-based representations (Paivio, 1986 and subsequent work, especially 2007), and by the context availability model, which attributes this advantage to the richer and more readily accessible contexts associated with concrete words (Schwanenfluegel anf Shoben, 1983; Schwanenfluegel, 1991). More recent accounts include grounding theories such as the situated conceptualization framework, which emphasizes the dynamic and context-dependent nature of meaning, especially for abstract concepts (Barsalou anf Wiemer-Hastings, 2005; Barsalou, 2008; Barsalou et al., 2018), and embodied theories of semantic representation, which argue that meaning–particularly for abstract terms–draws on sensorimotor and emotional experiences (Altarriba et al., 1999; Vigliocco anf Kita, 2007; Kousta anf Vigliocco, 2011, 2014).

The referents of abstract words are mental constructs. It is impossible to infer the meanings of nouns like *existence* (CPD mean concreteness – MC 1.8/5) or *guilt* (MC 2.1/5), or the meanings of verbs like *perceive* (MC 1.6/5) or *appreci-*

ate (MC 1.8/5), based on sensorimotor experience or pointing. Language mediation is necessary to understand their meanings. Although a wide range of theoretical frameworks have been proposed to explain how abstract concepts are mentally represented and processed (for a review, see Pecher, Boot, anf Van Dantzig, 2011), more recent accounts have highlighted the particular importance of linguistic and affective mechanisms in their grounding. Unlike concrete concepts, which are often rooted in perceptual and motor experiences, abstract concepts seem to lack a direct link to sensorimotor modalities. Consequently, researchers have argued that their meanings are primarily derived from patterns of co-occurrence in language, metaphorical mappings, and emotional valence. These contemporary perspectives (Hinojosa, Moreno, anf Ferré, 2020; Kousta et al., 2011; Vigliocco et al., 2011) propose that language provides a structured framework through which abstract concepts acquire meaning, while affective experiences contribute additional depth and specificity. In this view, linguistic distributional information and emotional resonance jointly compensate for the absence of direct experiential referents, enabling the cognitive system to represent and use abstract knowledge effectively. Based on our previous research with elementary and high school students and teachers in the project ”The Building Blocks of Croatian Mental Grammar: Constraints of Information Structure”, we hypothesize that these words are mainly encountered through the educational settings, and that their meanings are enhanced with RREs. Therefore, we hypothesize that they will be relatively rare and restricted in less educated and less read populations. Their subjective frequency varies considerably more than that of concrete words across populations because many speakers find it difficult to remember these words or their meanings.

In the present study, we do not seek to evaluate any single theoretical account outlined above; rather, our focus is on the scope and depth of abstract vocabulary within individuals across specific social contexts. Namely, we know that the comprehension of abstract words relies heavily on rich linguistic experience because, unlike concrete concepts, abstract concepts lack direct sensorimotor grounding and are instead primarily represented through language-based knowledge. For individuals with limited reading experience (LRE) or restricted vocabulary–often referred to as having low rich reading experience (RRE)–the linguistic input necessary to form these representations is significantly diminished.

Reading exposes individuals to a wide range of contexts in which abstract words are used, allowing them to build nuanced semantic networks and infer meaning through co-occurrence, syntactic patterns, and metaphorical extension. Without frequent exposure to such varied linguistic environments, individuals are less likely to encounter the subtle and diverse contexts that give abstract terms their meaning. As a result, they are less able to generate or recognize semantic associations involving abstract concepts.

Moreover, research in distributional semantics (e.g., Landauer anf Dumais, 1997) has shown that word meaning–particularly for abstract terms–is constructed through the accumulation of linguistic experiences across different contexts. A limited vocabulary and reduced exposure to text thus constrain the development of dense semantic neighborhoods, weakening the ability to form and retrieve semantic associations for abstract words. In essence, individuals with low RRE lack the linguistic scaffolding necessary for grounding abstract concepts. This not only hampers their comprehension but also impairs their ability to activate related concepts, infer meaning from context, and engage in flexible semantic reasoning–capabilities crucial for understanding and using abstract language.

Based on the findings that adult native speakers of the same language do not share the same mental grammar and the same vocabularies because they attend to different cues in the input (Dąbrowska 2012, 2014, 2015), we conducted a study to understand the factors preventing speakers from building plausible meanings of abstract verbs and their collocations.

3. The Study

3.1. Participants

Through November and December 2023, three pen-and-paper questionnaires based on minimal semantic pairs of abstract prefixed verbs and pseudowords were completed by two groups of participants (for recent pseudowords research see Ulicheva et al. 2021; Gatti et al, 2023). All participants gave informed consent for the research, anonymity was guaranteed, and basic demographic information (age, gender, educational attainment, and self-assessment of reading)

was collected.[3] The first group consisted of 30 students from the University of Zagreb's Faculty of Humanities and Social Sciences (further UNIS), average age 19.2 years, M/F = 8/22. They all completed at least 12 years of education and reported between 3 and 5 hours of reading at least 3-pages long texts per day. The second group consisted of 30 students from different vocational high schools in Zagreb (further HSS). Their average age was 17.1 years, M/F ratio 16/14. On average, they completed 8 years of elementary and 2 years of high school education prior to testing and were not exposed to RREs, because vocational schools in Croatia are known for students rarely interested in reading. On average, they reported less than an hour of reading activity per day. Although there is an age difference of 2 years between two groups, based on our previous research (Gnjidić, Keresteš, Peti-Stantić, 2021; Peti-Stantić, 2019; Peti-Stantić, 2020), we argue that the disparity in education and the amount and depth of daily reading is the decisive factor in the difference between the two mentioned populations, and not the age difference. Also, we acknowledge that categorizing participants into UNIS and HSS may oversimplify their complex educational paths, but the author's extensive experience working with high school and university students in researching their reading habits and their comprehension of complex text structures (Stantić and Peti-Stantić 2021 and several dozen workshops addressing reading comprehension that the author conducted over the past seven years), as well as their self-reported reading assessment in this study, support the notion that HSS vocational students lack RRE while UNIS students are exposed to more complex texts and vocabulary.

3.2. Materials

The first two questionnaires comprised related association-based experiments (one with pseudowords, the other with CPVs), while the third questionnaire consisted of a verbal object recall task (ORC). Associations were selected as a powerful tool for acquiring concepts and words as well as building an abstract vocabulary network. ORC was chosen to test metalinguistic awareness of dif-

[3] This work was conducted as an extension of the Croatian Science Foundation project HRZZ-IP-2016-06-1210, The Building Blocks of Croatian Mental Grammar: Constraints of Information Structure (MEGAHR; http://megahr.ffzg.unizg.hr/en/) awarded to Professor Anita Peti-Stantić. Thus, the ethical approval for the entire project applies to this work.

ferent collocations for closely semantically related verbs. All experiments were timed (2 minutes per stimulus).

Based on two CPV pairs chosen for the study, we created 6 pseudowords for the experiment 1, while we used CPVs for the experiment 2. The pseudowords were based on subword cues (Gatti et al., 2023) either on the prefixes they share with CPVs (*pri-* and *pro-*, with pseudowords *prilenati, proveliti*) or the base-verbs they share with CPVs (*pisati* 'to write', *dati* 'to give', *cijeniti* 'to appreciate', and *suditi* 'to judge', with pseudowords *zvocijeniti, blosuditi, klidati, dripisati*). Participants had to generate as many words as possible that they assumed would help identify the meaning of the six provided pseudowords and four CPVs. This test was administered first to mitigate the influence of memorization of entire CPVs on responses. Each pseudoword was provided on a separate piece of paper after the instructions had been read.

Focusing on identifying phonological or semantic links between the answers and given pseudowords, first experiment sought to determine the number and type of associative connections that participants could make with an unknown word meaning. The purpose of this experiment was to simulate conditions in which people either do not comprehend an abstract word or only understand it very loosely. We developed two predictions: 1. HSS students were expected to generate significantly fewer associations than UNIS, 2. root based pseudowords will result in more semantic associations as opposed to prefix based pseudowords, which were predicted to lead to more phonological associations.

As already stated, for the second and third experiment, two specific abstract verbal pairs were chosen for their semantic proximity, but distance in word-frequency, as well as somewhat salient collocational potential. These pairs were:

1. *pripisati* 'to attribute' and *pridati*, 'to add, to attribute',

2. *procijeniti* 'to appraise' and *prosuditi* 'to judge'.

The meanings of the verbs within the pairs are almost synonymous and interchangeable, but their subjective and objective frequencies, as well as collocational salience, were different. Differences in pair 1 objective corpus frequency were 6:1 and 8:1 for pair 2. The Sketch Engine data, however, showed an inverse relationship between corpus frequency and salience of their constructions.

Due to the close semantic proximity of stimuli, participants in the second and third experiment also received the PCVs on separate sheets of paper. Corresponding to the first experiment, in the second one, the participants were asked to provide as many words as possible which they thought to be semantically related to the prompt.

The third questionnaire consisted of the same CPVs, but the participants had to provide as many objects to these verbs as possible. The motivation for this task was to compare participants' abstract mental lexicons to collocations attested in Sketch Engine. The aim of the third experiment was twofold – to compare the number of produced unique objects for presented CPVs and to compare the number of Sketch Engine-aligned objects in two populations. Therefore, two metrics were considered in the coding scheme: 1) the number of unique words produced and 2) the number of words confirmed as collocations in Sketch Engine.

3.3. Hypothesis

We hypothesized that UNIS would generate more responses than HSS, both in the association task and in the ORC. We also assumed that the prevalent type of associations (phonological of semantic) each group generated indicates the difference in the density of their abstract mental lexicons. Those who rely more on phonological associations (PA) would be assumed to have developed less extensive monitoring process for abstract meanings (Borghi et al., 2021) compared to those who rely more on semantic associations (SA). This expectation comes from the notion that phonology serves as a stepping-stone and an intermediary for mapping the meanings of abstract words onto their corresponding lexical forms, which is consistent with the phonological cues bias confirmed in infant word learning (Culbertson anf Schuler, 2019) and the contrast between phonological and semantic cues in aphasic people (Meteyard anf Bose, 2018). For some speakers, this mapping process ends in phonology, frequently with incorrect conclusions.

We also predicted that UNIS would perform better on the unique object measure and the confirmed Sketch Engine collocations measure in the third experiment. This assumption is based on the well-established fact that a person's level of

comprehension and application of specific collocations is influenced by their exposure to abstract vocabulary and complex texts during their education.

3.4. Results

To determine whether the associations are phonologically or semantically driven, we developed a coding scheme. If the generated word was semantically connected with either the recognizable prefix or root, we coded it S, while obvious phonological similarity with the stimulus was coded P. The rest was classified as indeterminate (I). In accordance with that, we observed numerous phonological associations connected to recognizable prefixes such as *prileći, priložiti, prispojiti, pripisati, pridati, prilijepiti, proletjeti, propitati, provaliti, provjeriti, provoditi etc.* in both groups. However, semantic associations connected to phonologically recognizable, but semantically unrecognizable part of the word, such as *odrijemati, spavati, odmoriti, pasati, ljenčariti, zaspati, krevet, jastuk* for the first pair of pseudowords and *kazati, govoriti, početi govoriti, ispričati, objasniti* was insignificant. Furthermore, we observed both semantic and phonological associations in connection to pseudowords with recognizable semantic meaning. Semantic ones were *cijeniti, poštovati, procjena, osuditi, prosuditi, procijeniti*, produced by significant number of participants, while some produced small number of associations connected to the semantically unrecognizable prefix as well, such as *zvono, zvocanje, blokiranje, blud*. This demonstrates that the participants searched for semantic cues even in their absence.

As hypothesized, in a paired samples t-test HSS showed significantly fewer associations than UNIS, both phonological (HSM = 9.57, Std = 4.01 vs UNIM = 16.60, Std = 5.43), $t(58) = 5.70$, $p < 0.001$; and semantic (HSM = 11.27, Std = 4.25 vs UNIM = 25.47, Std = 7.28), $t(58) = 8.82$, $p < 0.001$. An exploratory analysis also revealed a significant difference between HSS (M = 11.47, Std = 5.83) and UNIS (M = 18.43, Std = 6.39) in the raw number of indeterminate associations, $t(58) = 4.41$, $p < 0.001$. However, an analysis of the ratios of indeterminate to differentiated associations, calculated at the level of participant, did not reveal significant differences between HSS (M = 0.58, Std = 0.31) and UNIS (M = 0.46, Std = 0.19), $t(58) = 1.82$, $p = 0.74$.

Examining differences within each population, patterns of data were largely as hypothesized. Paired-sample t-tests were run, and given the directional nature of these hypothesis, one-sided p-test values are reported (but see two-sided values in square brackets for transparency). In the HSS group, the number of PA (M = 3.13, Std = 1.04) significantly exceeded the number of SA (M = 0.27, Std = 0.45), t(29) = 13.46, p < 0.001 [0.001] for prefix based pseudowords. In the same group, for root based pseudowords, the number of PA (M = 6.43, Std = 3.41) was expectedly lower than the number of SA (M = 11.00, Std = 4.14), t(29) = -5.52, p < 0.001 [0.001].

This pattern replicated partially in the UNIS group. With prefix based pseudowords, the number of PA (M = 5.70, Std = 3.86) was not significantly different from the number of SA (M = 5.73, Std = 3.79), t(9) = -0.04, p = 0.49 [0.97]. However, in root based pseudowords the number of PA (M = 10.90, Std = 3.99) was, as expected, significantly lower than the number of SA (M = 19.73, Std = 7.24), t(29) = -6.05, p < 0.001 [0.001].

Two weeks after the first questionnaire, the same task was conducted with CPVs *pridati, pripisati, prosuditi* and *procijeniti*. Semantic transparency of the stimuli supported a prediction that all participants will generate more associations in this experiment than in the previous one.

Correcting for number of trials between first and second experiment (6 pseudowords in Experiment 1 vs 4 real verbs in Experiment 2), we observed a significant difference in association counts, with a significantly higher number of associations for CPVs (M = 5.82, Std = 2.46) compared to pseudowords (M = 5.24, Std = 2.22), t(59) = 2.05, p = 0.045.

As expected, HSS (M = 6.37, Std = 2.80) showed significantly fewer PA than UNIS (M =7.20, Std = 4.31), t(58) = -4.94, p = 0.19. The difference was more pronounced in SA, with HSS (M = 11.60, Std = 4.79) showing significantly fewer associations than UNIS (M = 21.37, Std = 7.85), t(58) = -5.82, p < 0.001. An exploratory analysis also revealed a significant difference between HSS (M = 3.93, Std = 3.38) and UNIS (M = 0.70, Std = 1.06) in the raw number of indeterminate associations, t(58) = 5.00, p < 0.001, albeit in this case the direction was different than in Experiment 1. An analysis of the ratios of indeterminate to differentiated

associations also revealed significant differences between HSS (M = 0.23, Std = 0.19) and UNIS (M = 0.03, Std = 0.05), t(58) = 5.47, p < 0.001.

In the third experiment, we observed a difference in the number of unique objects between HSS (M = 17.37, Std = 5.92) and UNIS (M = 25.87, Std = 6.02), t(58) = -5.52, p < 0.001. HSS group produced 998 objects in response to the same stimuli, compared to 1523 objects in the UNIS group. The same pattern was observed when only considering Sketch Engine-aligned objects as well.

This difference held and was significant testing within each pair of verbs and for both uniquely identified objects and Sketch Engine-aligned objects, see Table 1 for summary.

Table 1

	Group (N)	M + Stdev	t-value	p-value
Pair 1 - unique	HSS (30)	8.23 ± 3.90	-4.93	< .001
	UNIS (30)	13.27 ± 4.00		
Pair 1 - Sketch Engine	HSS (30)	7.50 ± 3.63	-5.37	< .001
	UNIS (30)	12.77 ± 3.96		
Pair 2 - unique	HSS (30)	9.13 ± 3.01	-4.71	< .001
	UNIS (30)	12.60 ± 2.67		
Pair 2 - Sketch Engine	HSS (30)	8.40 ± 2.69	-5.64	< .001
	UNIS (30)	12.13 ± 2.43		

A 2x2 ANOVA investigating the effects of sample group (university vs. high school) and word pair (*procijeniti* anf *prosuditi* vs. *pripisati* anf *pridati*) on the number of Sketch Engine-aligned objects was conducted. There was no interaction between sample group and word pair (F(1, 116) = 1.68, p = 0.20) nor main effect of word pair (F(1,116) = 0.51, p = 0.82). The main effect of sample group was significant (F(1, 116) = 57.88, p < 0.001), in line with the above.

Considering only Sketch Engine-aligned objects, we further broke this down to investigate differences within each pair. In the first pair (*pripisati* vs. *pridati*), a significant difference was observed between the number of Sketch Engine-aligned objects for *pridati* (M = 4.05, Std = 2.34) and *pripisati* (M = 6.08, Std = 2.64), t(59) = 13.41, p < 0.001. The same pattern replicated in the second pair (*procijeniti* vs. *prosuditi*), with a significant difference observed between the number of Sketch Engine-aligned objects for *prosuditi* (M = 3.27, Std = 1.54) and *procijeniti* (M = 7.00, Std = 2.03), t(59) = 16.44, p < 0.001.

4. Discussion and conclusions

Although the acquisition and processing of abstract words are influenced by a variety of linguistic, psychological, and social factors, this study looked only in how the social ones–educational attainment and exposure to RRE–impact the availability of the abstract vocabulary network, which is reflected in the capacity to form associations and recall objects. According to the lexical quality hypothesis (Perfetti anf Hart, 2002), the capacity to learn new words depends on developing high-quality lexical representations for many lexical entries. Given the extent to which abstract words depend on linguistic input for understanding, this is more significant for abstract than for concrete words. Orthography and phonology, as two out of three components of high-quality representation, get to be well-integrated and readily activated in processing in Croatian, because of transparent orthography. But, instead of focusing only on phonology in overcoming comprehension difficulties (Brennan anf Ireson, 1997), one should integrate the semantics, too. Our study showed that the sensitivity to semantic cues is enhanced in better educated individuals and populations exposed to RRE, while the less educated remain at the phonological level.

Despite being based on a subjective coding scheme created by the author, the distinction between phonological and semantic associations is supported by research on the differences between the two (e.g., De Deyne and Storms, 2008; Vivas et al., 2018; Haslett and Zhenguang, 2024; Hippolyte et al., 2025 and others). The author of this article is also aware that the only way to quantify theoretical cognitive models of lexical access and develop the interventions required to increase the abstract vocabulary in the educational system is to use multilayer architecture that considers both tiers of mental grammar (see Levy et al., 2021). Therefore, this study should be taken as a first experimental step in this direction.

The aim of the first two experiments was to determine whether differences in educational attainment and exposure to RRE impact availability and activation levels of abstract lexicon. In both experiments, there was a significant difference in the number of associations between the two populations. As expected, the difference was more noticeable with pseudowords than with real verbs. The differences between PA and SA in both experiments in the HSS group were

as predicted. Also, reliance of HSS on phonological cues even there where the semantic ones were offered, suggests that the semantic tier of the language architecture is insufficiently available to them. At the same time, UNIS generated nearly twice as many semantic associations as HSS, searching for semantic associations even in the absence of semantic cues, indicating a far higher degree of semantic tier availability. The two groups' indeterminate associations in relation to real words differed substantially, with UNIS producing significantly fewer indeterminate associations, indicating that recalling verbal meanings enabled production of a large number of semantically accurate associations or prevented speculation.

In the object recall task, the density and diversity of the collocational network was investigated. Here, the ability to recall objects required recognition of semantic and syntactic frames of complex prefixed verbs in addition to the specificity and salience of their collocations (Bolognesi et al., 2020). As a result, we interpret the difference in results between the populations as the difference in their metalinguistic awareness. Fewer unique Sketch Engine-aligned objects in HSS shows their limited capacity for producing accurate and diverse abstract word combinations.

Our results suggest that educational attainment and exposure to RRE impact the capacity to form associations and recall collocations within the abstract lexicon. Social implications of these findings are in line with the argument that the opportunities to encounter variety of words in rich and diverse linguistic environments are the key factor in cumulative nature of lexical learning (Hsiao anf Nation, 2018; Ricketts et al., 2021). We have shown that linguistic resources unavailable to certain social groups leave them without the integration of semantics in the meaningful way in the abstract lexical network. If false phonological associations do not get rectified, they do not support the cumulative nature of lexical learning, particularly for abstract words. Repairs can come in different forms, explicitly or implicitly. Sufficiently rich reading experience might be the key for developing sensitivity to all types of meaningful cues and their integration (Ulicheva et al., 2021).

Since the findings suggest that populations with lower educational attainment and RRE were unable to create high-quality lexical representations by integrating information from the phonological and semantic tiers (Jackendoff, 2002),

we not only found that the populations differ in size and depth of their abstract mental lexicons, but also that there is a parallelism between the types of cues available to a given population and the degree of integration and complexity of the language architecture they are metalinguistically aware of.

Finally, one reviewer points out that there might be additional explanations for the observed effects. Undoubtedly, students who do better on language and reading fluency tests may also do better on word association tasks. Additionally, HSS students may not have been as familiar with the selected pairs of CPVs as UNIS students. However, educational systems should aim to increase the number of students who achieve these competencies because it has been demonstrated that students who achieve higher levels of language and reading fluency have a deeper and broader understanding of the meanings of individual words and their collocations, especially the abstract ones. Future research on social mediation and distribution of abstract vocabulary should certainly use larger, more varied samples, include objective assessments of reading exposure and cognitive capacity, and most likely create better methods for lexical association coding and analysis. I appreciate two anonymous reviewers for bringing this to my attention.

Bibliography

Altarriba, Jeanette; bauer, lisa M.; Benvenuto, Claudia. 1999. Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, anf Computers* 31/4. 578–602.

Barber, Helen A.; Otten, Lianne J.; Kousta, Stavroula-Thaleia; Vigliocco, Gabriella. 2013. Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and Language* 125/1. 47–53. https://doi.org/10.1016/j.bandl.2013.01.005.

Barsalou, Lawrence W. 2008. Grounded cognition."*Annual Review of Psychology* 59. 617–645.

Barsalou, Lawrence W.; Dutriaux, Lorena; Scheepers, Christop. 2018. Moving beyond the distinction between concrete and abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences* 373. 20170144. https://doi.org/10.1098/rstb.2017.0144.

Barsalou, Lawrence W.; Wiemer-Hastings, Katja. 2005. Situating abstract concepts.

Grounding cognition: The role of perception and action in memory, language, and thought. Eds. Pecher, Diane; Zwaan, Rolf A. Cambridge University Press. 129–163.

Binder, Jeffrey R.; Westbury, Chris F.; McKiernan, Kelly A.; Possing, Eliza T.; Medler, David A. 2005. Distinct Brain Systems for Processing Concrete and Abstract Concepts. *Journal of Cognitive Neuroscience* 17/6. 905–917.

Bolognesi, Marianna; Burgers, Christian; Caselli, Tommaso. 2020. On abstraction: decoupling conceptual concreteness and categorical specificity. *Cognitive Processing* 21/3. 365–381. https://doi.org/10.1007/s10339-020-00965-9.

Borghi, Anna M.; Binkofski, Ferdinand. 2014. *Words as Social Tools: An Embodied View on Abstract Concepts*. Springer. New York. https://doi.org/10.1007/978-1-4614-9539-0.

Borghi, Anna M.; Fini, Claudia; Tummolini, Luca. 2021. Abstract concepts and metacognition: Searching for meaning in self and others. *Handbook of embodied psychology: Thinking, feeling, and acting*. Eds. Robinson, Michael D.; Thomas, Laura E. Springer Nature Switzerland AG. 197–220. https://doi.org/10.1007/978-3-030-78471-3_9.

Brennan, Fiona; Ireson, Judith. 1997. Training phonological awareness: A study to evaluate the effects of a program of metalinguistic games in kindergarten. *Reading and Writing* 9. 241–263.

Castles, Anne; Rastle, Kathleen; Nation, Kate. 2018. Ending the reading wars: Reading acquisition from novice to expert. *Psychological science in the public interest* 19/1. 5–51. https://doi.org/10.1177/1529100618772271.

Cevoli, Barbara, Watkins, Calum; Rastle, Kathleen. 2022. Prediction as a basis for skilled reading: insights from modern language models. *Royal Society Open Science* 9/6. 211837. https://doi.org/10.1098/rsos.211837.

Croft, William. 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. The University of Chicago Press.

Culbertson, Jennifer; Schuler, Katherine. 2019. Artificial language learning in children. *Annual Review of Linguistics* 5. 353–373. https://doi.org/10.1146/annurev-linguistics-011718-012329.

Dąbrowska, Ewa. 2012. Different speakers, Different grammar: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2/3. 219–253. https://doi.org/10.1075/lab.2.3.01dab.

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

Dąbrowska, Ewa. 2015. Individual differences in grammatical knowledge. *Handbook of Cognitive Linguistics*. Eds. Dąbrowska, Ewa; Divjak, Dagmar. De Gruyter Mouton. Berlin. 650–668. https://doi.org/10.1515/9783110292022-033.

De Deyne, Simon; Storms, Gert. 2008. Word associations: Network and semantic properties. *Behavior Research Methods* 40(1). 213–231.

De Groot, Annette M. B. 1989. Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15/5. 824–845.

De Groot, Annette M. B.; Keijzer, Rina. 2000. What Is Hard to Learn Is Easy to Forget: The Roles of Word Concreteness, Cognate Status, and Word Frequency in Foreign-Language Vocabulary Learning and Forgetting. *Language Learning*, 50/1. 1–56.

Fliessbach, Klaus; Weis, Susanne; Klaver, Peter; Elger, Christian E.; Weber, Bernd. 2006. The effect of word concreteness on recognition memory. *NeuroImage* 32/3. 1413–1421.

Gatti, Daniele; Marelli, Marco; Rinaldi, Luca. 2023. Out-of-vocabulary but not meaningless: Evidence for semantic-priming effects in pseudoword processing. *Journal of Experimental Psychology: General* 152/3. 851. https://psycnet.apa.org/doi/10.1037/xge0001304.

Gentner, Dedre. 1982. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *BBN report; no. 4854*.

Gleitman, Lila R.; Cassidy, Kimberly; Nappa, Rebecca; Papafragou, Anna; Trueswell, John C. 2005. Hard words. *Language learning and development* 1/1. 23–64.

Gnjidić, Vedrana; Keresteš, Gordana; Peti-Stantić, Anita. 2021. Mjerenje širine i dubine akademskog vokabulara učenika viših razreda osnovne škole. *Croatian Journal of Education = Hrvatski časopis za odgoj i obrazovanje* 23/4. 1043–1085. https://doi.org/10.15516/cje.v23i4.4145.

Haslett, David A.; Zhenguang G. Cai. 2024. Wayward associations: When and why people think of similar-sounding words. *Journal of Memory and Language* 138. 104537. https://doi.org/10.1016/j.jml.2024.104537.

Hinojosa, Juan A.; Moreno, Emilia M.; Ferré, Pilar. 2020. On the limits of affective neurolinguistics: a "universe" that quickly expands. *Language, Cognition and Neuroscience* 35/7. 813–839. https://doi.org/10.1080/23273798.2020.1761988.

Hsiao, Yaling; Nation, Kate. 2018. Semantic diversity, frequency and the development of lexical quality in children's word reading. *Journal of Memory and Language* 103. 114-126. https://doi.org/10.1016/j.jml.2018.08.005.

Hippolyte, Amélie; Ribeiro, Nadine; Ibernon, Laurent; Marec-Breton, Nathalie; Declercq, Christophe. 2025. Semantic and phonological associative norms of 145 French words for children aged 5–9 years: Free associations and forced-choice task. *First Language* 45/2. 177–195. https://doi.org/10.1177/01427237251313567.

Jackendoff, Ray S. 1994. *Patterns in the mind: Language and human nature*. Basic Books.

Jackendoff, Ray S. 2002. *Foundations of language*. Oxford University Press. Oxford, UK.

Kounios, John; Holcomb, Phillip J. 1994. Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20/4. 804–823

Kousta, Stavroula-Thaleia; Vigliocco, Gabriella. 2011. The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General* 140/1. 14–34. https://doi.org/10.1037/a0021446.

Kousta, Stavroula-Thaleia; Vigliocco, Gabriella; Vinson, David P.; Andrews, Mark; Del Campo, Elena. 2011. The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General* 140/1. 14–34. https://doi.org/10.1037/a0021446.

Landauer, Thomas K.; Dumais, Susan T. 1997. A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review* 104/2. 211–240.

Langacker, Ronald W. 1987. Nouns and Verbs. *Language* 63/1. 53–94.

Lee, Chia-lin; Federmeier, Kara D. 2008. To watch, to see, and to differ: An event-related potential study of concreteness effects as a function of word class and lexical ambiguity. *Brain and Language* 104/2. 145–158.

Levy, Orr; Kenett, Yoed N.; Oxenberg, Orr; Castro, Nichol; De Deyne, Simon; Vitevitch, Michael S.; Havlin, Shlomo. 2021. Unveiling the nature of interaction between semantics and phonology in lexical access based on multilayer networks. *Sci Rep* 11. 14479. https://doi.org/10.1038/s41598-021-93925-y.

Ma, Weiyi; Michnick Golinkoff; Roberta, Hirsh-Pasek; Kathy, McDonough, Colleen; Tardif, Twila. 2009. Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language* 36/2. 405–423.

Meteyard, Louisa; Bose, Arpita. 2018. What does a cue do? Comparing phonological and semantic cues for picture naming in aphasia. *Journal of Speech, Language, and Hearing Research* 61/3. 658–674. https://doi.org/10.1044/2017_JSLHR-L-17-0214.

Nelson, Katherine; Hampson, June; Shaw, Lea K 1993. Nouns in early lexicons: evidence, explanations and implications. *Journal of Child Language* 20/1. 61–84.

Paivio, Alan. 1986. *Mental representations: A dual coding approach*. Oxford University Press.

Paivio, Alan. 2007. *Mind and its evolution: A dual coding theoretical approach.* Psychology Press. New York, NY.

Pecher, Diane; Boot, Inge; Van Dantzig, Saskia. 2011. Abstract concepts: Sensory-motor grounding, metaphors, and beyond. *The Psychology of Learning and Motivation*.

Ed. Ross, Brian H. Vol. 54. Academic Press. 217–248. https://doi.org/10.1016/B978-0-12-385527-5.00007-3.

Perfetti, Charles A.; Hart, Laurie. 2002. The lexical quality hypothesis. *Precursors of functional literacy*. Eds. Verhoeven, Ludo; Elbro, Carsten; Reitsma, Pieter. John Benjamins Publishing Company. 189–213.

Peti-Stantić, Anita. 2019. *Čitanjem do (spo)razumijevanja: od čitalačke pismenosti do čitateljske sposobnosti.* Naklada Ljevak. Zagreb. 2019.

Peti-Stantić, Anita. 2020. Data to Help us Form Images: The Abstract and Non- Imageable. *NooJ 2020 Conference*. (plenarno izlaganje)

Peti-Stantić, Anita; Anđel, Maja; Gnjidić, Vedrana; Keresteš, Gordana; Ljubešić, Nikola; Masnikosa, Irina; Tonković, Mirjana; Tušek, Jelena; Stanojević, Mateusz -Milan. 2021. The Croatian psycholinguistic database: Estimates for 6000 nouns, verbs, adjectives and adverbs. *Behavior Research Methods* 53/4. 1799–1816. https://doi.org/10.3758/s13428-020-01533-x.

Ricketts, Jessie; Dawson, Nuala; Davies, Robert. 2021. The hidden depths of new word knowledge: Using graded measures of orthographic and semantic learning to measure vocabulary acquisition. *Learning and Instruction* 74. 101468. https://doi.org/10.1016/j.learninstruc.2021.101468.

Romani, Cristina; McAlpine, Sarah; Martin, Randi C. 2008. Concreteness Effects in Different Tasks: Implications for Models of Short-Term Memory. *Quarterly Journal of Experimental Psychology* 61/2. 292–323.

Schwanenflugel, Paula J. 1991. Why are abstract concepts hard to understand? *The psychology of word meanings*. Ed. Schwanenflugel, Paula J. Lawrence Erlbaum Associates. 223–250.

Schwanenflugel, Paula J.; Akin, Ceylan; Luh, Wen-Ming. 1992. Context availability and the recall of abstract and concrete words. *Memory anf Cognition* 20/1. 96–104.

Schwanenflugel, Paula J.; Shoben, Edward J. 1983. Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9/1. 82–102.

Simonsen, Hanne Gram; Lind, Marianne; Hansen, Pernille; Holm, Else; Mevik, Bjørn-Helge. 2013. Imageability of Norwegian nouns, verbs and adjectives in a cross-linguistic perspective. *Clinical Linguistics anf Phonetics* 27/6-7. 435–446. https://doi.org/10.3109/02699206.2012.752527.

Stantić, Vedrana; Peti-Stantić, Anita. 2021. *Znati(želja): zašto mladi trebaju čitati popularnoznanstvene tekstove, i to odmah?*. Naklada Ljevak. Zagreb.

Tomasello, Michael; Akhtar, Nameera; Dodson, Katherine; Rekau, Linda. 1997. Differential productivity in young children's use of nouns and verbs. *Journal of Child Language* 24/2. 373–387. https://doi.org/10.1017/S0305000997003085.

Ulicheva, Anastasia; Coltheart, Max, Grosseck, Otto; Rastle, Kathleen. 2021. Are people consistent when reading nonwords aloud on different occasions?. *Psychonomic Bulletin anf Review* 28/5. 1679–1687. https://doi.org/10.3758/s13423-021-01925-w.

Vigliocco, Gabriella; Kita, Sotaro. 2006. Language-specific properties of the lexicon: Implications for learning and processing. *Language and Cognitive Processes* 21/7-8. 790–816.

Vigliocco, Gabriella; Vinson, David P.; Druks, Judit; Barber, Helen; Cappa, Stefano F. 2011. Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience anf Biobehavioral Reviews* 35/3. 407–426. https://doi.org/10.1016/j.neubiorev.2010.04.007.

Vivas, Laura; Manoiloff, Laura; García, Adolfo M.; Lizarralde, Federico; Vivas, Julio. 2018. Core Semantic Links or Lexical Associations: Assessing the Nature of Responses in Word Association Tasks. *Journal of Psycholinguistic Research* 48. 243–256. https://doi.org/10.1007/s10936-018-9601-8

West, Wally C.; Holcomb, Phillip J. 2000. Imaginal, Semantic, and Surface-Level Processing of Concrete and Abstract Words: An Electrophysiological Investigation. *Journal of Cognitive Neuroscience* 12/6. 1024–1037.

Društvena konstrukcija apstraktnog leksikonaSažetak

Apstraktna značenja riječi, koja su u potpunosti mentalni konstrukti, ovise o posredovanju jezikom te ih se u tom kontekstu uči, prisjeća i koristi. Nedavna otkrića pokazala su da odrasli izvorni govornici istog jezika ne dijele istu mentalnu gramatiku i vokabular jer obraćaju pažnju na različite ključeve prilikom usvajanja jezika. Stoga smo proveli ispitivanje koje istražuje ulogu obrazovnih postignuća i izloženosti bogatim čitalačkim iskustvima (engl. *rich reading experiences*, RRE) kao društvenih čimbenika koji utječu na sposobnost stvaranja asocijacija i dosjećanja kolokacija unutar apstraktnog leksikona. Primarni ciljevi studije bili su: 1. utvrđivanje odnosa između društvenih čimbenika i sposobnosti stvaranja asocijacija i dosjećanja kolokacija unutar apstraktnog leksikona i, 2. razumijevanje paralelizma između društvenih čimbenika koji utječu na potencijal za izgradnju apstraktnog leksikona i tipova ključeva kodiranih u jezičnoj arhitekturi. Utvrdili smo da se ispitane populacije, srednjoškolska i studentska, razlikuju u širini i dubini svojih apstraktnih mentalnih leksikona, ali i to da se niže obrazovana populacija koja manje čita u većoj mjeri oslanja na fonološke, dok se više obrazovana populacija koja više čita u većoj mjeri oslanja na semantičke ključeve.

Ključne riječi: apstraktni leksikon, gustoća mreže, jezik kao društveni alat, obrazovanje,

bogato čitalačko iskustvo
Key words: abstract lexicon, network density, language as a social tool, education, rich reading experience