THE REPRESENTATION OF CONCRETE VERSUS ABSTRACT WORDS: AN EYE-TRACKING STUDY

Mahsa Morid, Ed Griffiths, Shuzhen Wang, and Tania S. Zamuner
University of Ottawa

Introduction

A word is considered concrete when it has available sensory referents and an easily accessible mental image. For example, the word 'bread' describes an object which can be directly tasted, smelled and touched. Compared to concrete words, abstract words (e.g., 'justice') lack the same quantity and/or types of direct sensory referents. Some researchers claim that the processing differences between concrete and abstract words are mainly due to the quantitative differences between the amount (and partly the type) of available information for these two categories of words (Schwanenflugel and Stowe 1989). On the contrary, other researchers have highlighted the qualitative differences between these two types of words. In this view, the claim is that concrete words have connections to other words in the mental lexicon that are "categorical" and mainly based on semantic similarity, while abstract words are organized in an associative network whose connections are primarily based on their association with other words (Crutch 2006). One of the first eye-tracking studies investigating qualitative/associative differences between concrete and abstract words on healthy participants was undertaken by Duñabeitia et al. (2009) in Spanish. Using the Visual World Paradigm (VWP) (Cooper 1974, Tanenhaus et al. 1995), Duñabeitia et al. found that when participants heard a word, they fixated on competitor images that represented an association with the target word. More importantly, this shift of visual attention was quicker and greater when the target-competitor pairs were abstract vs. concrete. They concluded that abstract words have easier access to their respective associated concepts than concrete words.

Given the importance of the organization and processing of concrete versus abstract words for language theories, we sought to replicate and extend the study by Duñabeitia et al. into English, to further investigate whether the difference between concrete and abstract words is rooted in the words' qualitative differences. Using a VWP paradigm, English native speakers heard a word while they saw four images (one target and three distractor images). Importantly, the audio stimuli on different trials varied in three ways with respect to the target image: Identical trials, Associated trials or Unrelated trials. Half of the trials had concrete objects as the associated audio (e.g., target image 'fish' paired with either audio 'fish' (identical), audio 'pond' (associated), audio 'tree' (unrelated) and half of the trials had abstract words as the associated audio (e.g., target image 'nose' paired with either audio 'nose' (identical), audio 'smell' (associated), audio 'moment' (unrelated). Each participant was exposed to only one of those three trials per condition (total of 30 displays). The critical trials for our analyses were the associated trials that were either concrete or abstract. Based on the association hypothesis, we hypothesized that the fixations to the

Actes du congrès annuel de l'Association canadienne de linguistique 2020.
Proceedings of the 2020 annual conference of the Canadian Linguistic Association.

target image should be faster and more frequent when the associated relationship was abstract (e.g., target 'nose', audio 'smell') vs concrete (e.g., target 'fish', audio 'pond'). The results showed that participants looked less frequently to the image of a fish when they heard 'pond', and in comparison looked more to the image of a nose when they heard 'smell'. Therefore, the results of the current study supported an associative network for abstract words and concepts.

The remainder of this paper is organized as follows: Section 1 focuses on a brief review of the previous literature that is important for understanding the purpose of the current study. In Section 2, the methodological details of the study are discussed, followed by the presentation of the results in Section 3. Finally, discussion of the results and the conclusion is provided in Section 4.

1. Background

As stated, concrete and abstract words have some dissimilarities, which in turn result in processing differences. The *concreteness effect* refers to the observation that concrete words are processed faster and more accurately than abstract words (Holcomb et al. 1999). Moreover, it has been shown that the time to comprehend a sentence is shorter when the sentence includes concrete words (West and Holcomb 2000), while sentences constructed of abstract words generally take longer to read (Schwanenflugel and Stowe 1989). There are two major theoretical accounts concerning the source of the concreteness effect: *dual-coding theory* and the *context-availability model*.

According to the dual-coding theory (Paivio 1986, 1991), the type of available information for concrete words is different from that of abstract words. Concrete words have access to information in both the 'verbal linguistic' system and 'nonverbal imagistic' system, while abstract words only have access to information stored in the verbal linguistic system. The processing advantage of concrete words thus arises from their availability to multiple sources of information. The context-availability model (Bransford and McCarrell 1974, Kieras 1978), on the other hand, argues that the quality of available information, but not the type of information, is different for concrete and abstract words. In other words, the advantage of concrete words over abstract words stems from the greater contextual connections in semantic memory for concrete words compared to abstract words.

Based on a framework which assumes 'qualitative' difference for abstract and concrete words, abstract words are organized in a neural network where the connection between them is based on *semantic association*. In contrast, concrete words are connected to each other based on *semantic similarity* (where, for example, the words 'theft' and 'punishment' have semantic associations, because theft will result in punishment, whereas the words 'theft' and 'burglary' are semantically similar.). The prediction of this theory is that once an abstract word is activated, the words that are associatively connected to that word will be co-activated. In turn, the activation of a concrete word will primarily co-activate the words that are semantically similar to that word. It is important to state that this does not mean that concrete words will not co-activate the words that are associatively connected to them, but that the activation of such words is slower compared to co-activation of abstract associates.

To empirically investigate this claim, Duñabeitia et al. (2009) designed a study using the Visual World Paradigm (VWP). In their study, participants saw four different pictures on the screen (one target word and three distractors) while they heard a word via headphones. In one condition, namely Abstract-Association, they heard an *abstract* word which was associated to the target picture (e.g., they heard the word 'smell', which is an abstract word, and they saw a picture of a nose). In another condition, i.e., Concrete-Association, participants heard a *concrete* word which was associated to the target word (e.g., they heard the word 'crib', which is a concrete word, and they saw a picture of a baby.). They hypothesized that if abstract words are primarily connected to their associates, then the proportion of looks toward the target picture should be higher (and the speed was likely to be faster) in the Abstract-Association condition, compared to Concrete-Association condition. The results of their study met their expectations. Their study was the first to investigate this issue in healthy adult participants. The current study extended the study by Duñabeitia et al. from Spanish to English.

2. Methods

2.1 Participants

Twenty-seven undergraduate students at the University of Ottawa were recruited for the experiment and participated in exchange for partial course credit. They were all native English speakers and reported normal or corrected-to-normal vision, normal hearing and no language impairment history.

2.2 Material Preparation

2.2.1 Word stimuli

In order to prepare word stimuli, we started by translating the words from Duñabeitia et al. (2009) from Spanish to English. From that list we removed some words for various reasons, for example, if a word had a morphologically complex translation in English such as 'lifeguard' or 'tablecloth'. We also removed words which had been judged abstract in Spanish but where the translation was concrete such as 'floor' or 'tool'. We also made sure that all the unrelated distractor words in the abstract conditions were abstract words and in the concrete conditions were concrete. We then checked the frequency of the words using log-transformed HAL frequency norms (Log_Freq_HAL) (Balota et al. 2007). This frequency tool consists of approximately 131 million words gathered across 3000 Usenet newsgroups in February 1995. The online tool that we used was the "English Lexicon Project Website", developed by Washington University. We then calculated mean and standard deviations of frequencies for all words and replaced the words that fell outside of mean +/-2SD. At this point, we had 17 sets of abstract and 21 sets of concrete stimuli (a total of 114 words).

As the next step toward norming our stimuli, we created an online survey to make sure that the degree of association between all the word pairs was judged to be similar. In

this survey, participants were told to rate (on a scale from 1 to 7) to what extent they considered the word pairs to be associated with each other. Responses were received from 5 participants. Despite the small scale of the survey, it gave us useful perspective on the selection of the final list of stimuli (e.g., we excluded the pair "foot-rope" in unrelated condition which received a high association rating of 3.5/7). The exclusions based on the results of this survey enabled us to come up with the final lists of 30 target words, 30 associated words (the critical condition) and 30 distractors. Table 1 represents the average-over-condition frequency and word length for the final stimuli.

Table 1. Average word frequency and word length per conditions.

	AI	**AA**	**AU**	**CI**	**CA**	**CU**
Word Frequency	3.37	2.89	2.64	3.34	2.92	2.83
Length (number of letters)	5.06	6.73	7.26	5.30	4.40	5.53

Note: AI: Abstract Identity; AA: Abstract Association; AU: Abstract Unrelated; CI: Concrete Identity; CA: Concrete Association; CU: Concrete Unrelated

2.2.2 Picture stimuli

We chose our stimuli pictures from the Multipic databank of a "750 publicly available color drawings representing common concrete concepts created by the same artist, standardized for name agreement and visual complexity in several European languages" (Duñabeitia et al. 2018: 809). These pictures were not used in the original Spanish study but were developed by its lead author. They were selected by us for the potential recognition advantage of colour images, along with norming data and multilingual standardization provided. In this database, "the modal name percentage" is provided, i.e., the percentage of 100 English native speaker participants who named the picture by its most commonly given name. When selecting our pictures, we made sure to select pictures with a high modal name percentage (above 75%).

The database accompanying the Multipic pictures also contained a visual complexity score based on the average rating for the complexity of the image given by 600 participants surveyed by Duñabeitia et al. (2018). We performed a t-test to compare the visual complexity of the target object for abstract and concrete stimuli. The result of the t-test revealed no significant difference between the visual complexity of the pictures for these two conditions: $t(14) = .55$, $p=.58$. We then performed another t-test to compare the average visual complexity in all the distractors in the abstract condition with the visual complexity of all the distractors in the concrete condition; again no significant difference was found: $t(14)= .94$, $p= .3$. With these measures taken, we made sure that, if any difference is observed between two experimental conditions, it will not be related to any stimuli-biased differences.

When choosing distractor images, we also made sure that there was no semantic or word-initial phonological overlap with the target, associated and unrelated words.

2.2.3 Audio stimuli

The audio stimuli consisted of a single word in each condition (e.g. 'lamp' for identity, 'electricity' for association, and 'youth' for the unrelated condition). The total of 90 words was recorded using Audacity software at a normal speaking rate by a female native speaker of Canadian English, using an ICICLE microphone. The intensity of the recordings was adjusted to an average of 70dB. To ensure clarity of recording each word was recorded three times. Then, based on a native speaker judgment, the best exemplar for each word was selected as the final audio file and edited so that the audio file began at the onset of the word. An example of stimuli is shown in Figure 1.

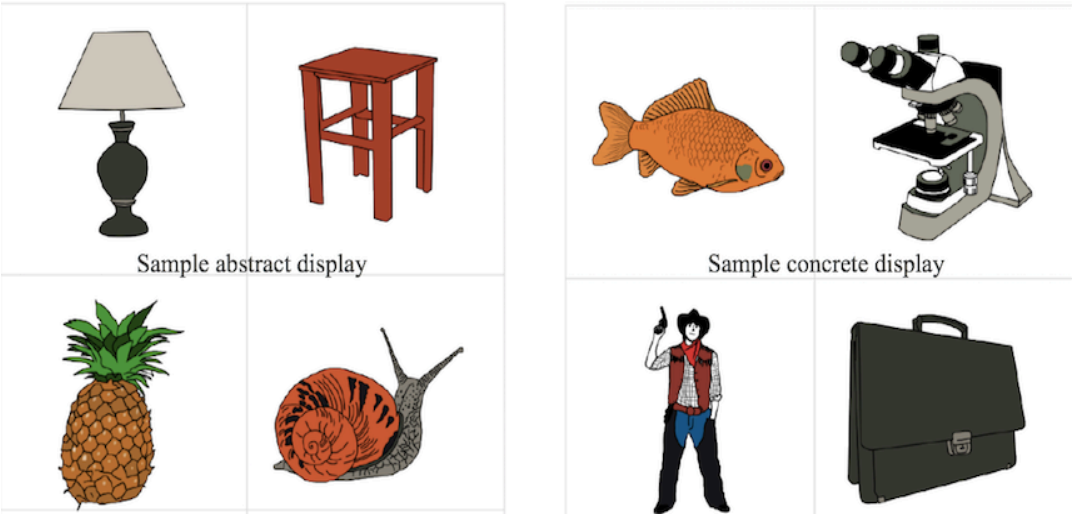

Figure 1. Example of abstract (left panel) and concrete (right panel) displays. In abstract display the target picture is *lamp* and the audio stimuli is *lamp* for Identity, *electricity* for Association, and *youth* for Unrelated conditions. In concrete display the target picture is *fish* and the audio stimuli is *fish* for Identity, *pond* for Association, and *tree* for Unrelated conditions.

2.3 Apparatus

The materials were presented using Experiment Builder software (SR Research, Ottawa) installed on an Apple Mac Mini. Participants had their dominant eye established using the Miles Test. Eye movements of the dominant eye were recorded with an EyeLink 1000 Eye Tracker (SR research, Ottawa), with chin rest, at a sampling rate of 500 Hz. Participants were seated in a chair 500-600mm from the camera. For each participant, calibration and validation were performed, based on a 5-point calibration grid in the form of a 'plus sign' at the beginning of the experiment. Drift correction was repeated between each trial.

2.4 Procedure

Participants were told that they would see four images and hear a word through their headphones. They were told that they should click on a picture if they considered that it matched the audio stimulus in some way. They were explicitly told that some of the audio words might not have a connection with any of the pictures. At the start of each trial, a fixation point was presented in the center of the screen. When the participant's gaze was steady on the centre point, the experimenter pressed the button to display the trial. The audio stimulus was presented 1000ms after the presentation of the pictures. The four pictures disappeared either 5000ms after the presentation of the audio or when the participant clicked on one of the pictures. The fixation point for the next trial then appeared.

Three blocks of pictures were created, and the same group of pictures was only presented once in each block. The trials were randomized and rotated within each block to make six lists in total. We balanced the six lists when planning the experiment. Participants were first presented with two practice trials and were asked whether they completely understood the instructions. The two practice displays were presented with all three audio stimuli: identity, associated and unrelated words sequentially. After the practice trials, the remaining trials were presented in a pseudorandom order under the constraint that the same group of pictures was only presented once under one condition within a block. There was an optional break time between each block. The locations of the four images within the visual display were randomized for each trial.

3. Results

Data were extracted using DataViewer software in 20 ms time bins (SR Research, Ottawa). Eye data from 26 participants were included in the analysis (one participant's eye data was lost due to equipment error). Data were first filtered so that only 'correct trials' according to click data were analyzed, i.e., trials where there was a mouse-click in the identity or associated condition, or no mouse-click in the unrelated condition. Participants' overall accuracy rate was 92.86%; a total of 167 out of 2340 trials were excluded. Analyses began 200 ms after the onset of the target, as this is the time it typically takes for adults to program an eye movement (Matin et al. 1993, Salverda et al. 2014). An inspection of the fixation data indicated that looks to the target peaked at 1000 ms after the target word onset, resulting in a window of analysis spanning from 200 to 1000 ms from the target word onset. The number of time bins was different from trial to trial since the trial ended after the mouse-click by the participant.

The data in the 200 to 1000 ms window were examined with a growth curve analysis (GCA) (Mirman et al. 2008), which is a multilevel regression technique allowing for the analysis of longitudinal data such as time course data. This technique allows for the calculation of differences in time, as well as the steepness of a looking curve (i.e., the rapidity of looking to the target picture). The analysis was done in R (R Core Team) using the lmer() function from the lme4 package (Bates et al. 2015).

The purpose of this analysis was to examine whether there was difference in looking behaviour when participants heard an abstract or a concrete word associated with

the target picture. The initial inspection of data showed that the overall looks in the Abstract-Association condition were higher compared to the Concrete-Association condition. Importantly, this difference did not seem to be present between Abstract-Identity and Concrete-Identity conditions. Figure 2 shows the proportion of looks in different conditions.

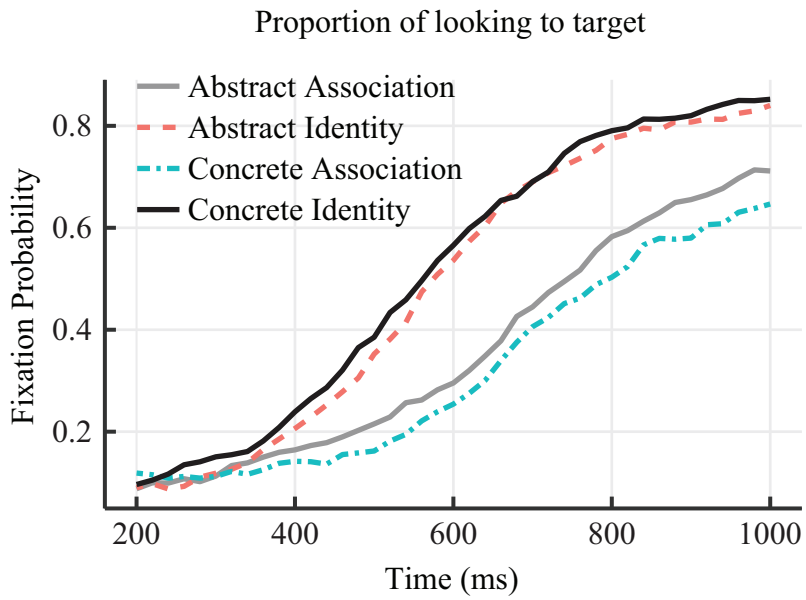

Figure 2. Proportion of looks to the target image in Identity and Association conditions, from 200 ms after auditory word onset.

To examine this difference statistically, a model was fitted with condition (abstract versus concrete) as the within-subject variable. Since condition was manipulated within participant, we included random effects of condition by-subject. Moreover, the model also included a main effect of Time (captured by orthogonal polynomials) and a condition by time interaction. Based on the shape of the looking data, a three-order (cubic) orthogonal polynomial model was considered to be the best fit for data. Table 2 shows the result of the statistical analysis. There was a significant effect of Condition on the intercept term, indicating lower overall target fixation proportions for the Concrete-Association relative to the Abstract-Association (*Estimate* = -4.04, *SE* = 0.01, p = 0.01). There were no significant effects on any orthogonal terms. This indicates that participants looked less to the image of a 'fish' when they heard 'pond', and in comparison looked more to the image of a 'nose' when they heard 'smell'. However, there was no difference in the speed of looking to the target images in the different conditions.

Table 2. GCA results for Association condition. GCA for the effect of abstractness (i.e., abstract association vs. concrete association) on looking data for a window of analysis 200-1000 ms after the auditory target word onset. Values represent parameter estimates.

Predictor	Estimate	*SE*	*t*	*p*
Intercept	0.45	0.01	24.00	<.001
Linear Term	1.85	0.10	18.48	<.001
Quadratic Term	-0.14	0.07	-1.81	0.07
Cubic Term	-0.33	0.06	-5.41	<.001
Condition (Concrete)	-0.04	0.01	-2.52	0.01
Condition* Linear Term	-0.13	0.13	-1.04	0.30
Condition* Quadratic Term	0.17	0.09	1.74	0.09
Condition* Cubic Term	0.02	0.07	0.26	0.79

Note. SE = standard error. GCA structure in R: lmer(PercentSampleCount1~ Condition * (Linear + Quadratic + Cubic) + (Linear + Quadratic + Cubic | Subject) + (Linear + Quadratic + Cubic | Subject: Condition).

A second model was fitted to investigate whether the same effect of Condition was present for Abstract Identity versus Concrete Identity (notice that this comparison is for control purpose). The fixed and random effects were similar to previous model. Here the shape of data indicated that a two-order (quadratic) model best described the data. As expected, the effect of condition was not significant, indicating that the effect found in Association Condition is not due to a stimuli-based bias. Table 3 shows the details of this analysis. Figure 3 shows the model prediction graph for two models.

Table 3. GCA results for Identity. GCA for the effect of abstractness (i.e., abstract identify vs. concrete identity) on looking data for a window of analysis 200-1000 ms after the auditory target word onset. Values represent parameter estimates.

Predictor	Estimate	*SE*	*t*	*P*
Intercept	0.59	0.02	29.99	<.001
Linear Term	2.08	0.08	24.19	<.001
Quadratic Term	-0.54	0.07	-7.69	<.001
Condition (Concrete)	0.01	0.01	1.46	0.15
Condition* Linear Term	-0.06	0.07	-0.92	0.36
Condition* Quadratic Term	-0.03	0.07	-0.52	0.60

Note. SE = standard error. GCA structure in R: lmer(PercentSampleCount1~ Condition * (Linear + Quadratic) + (Linear + Quadratic | Subject) + (Linear + Quadratic | Subject: Condition) .

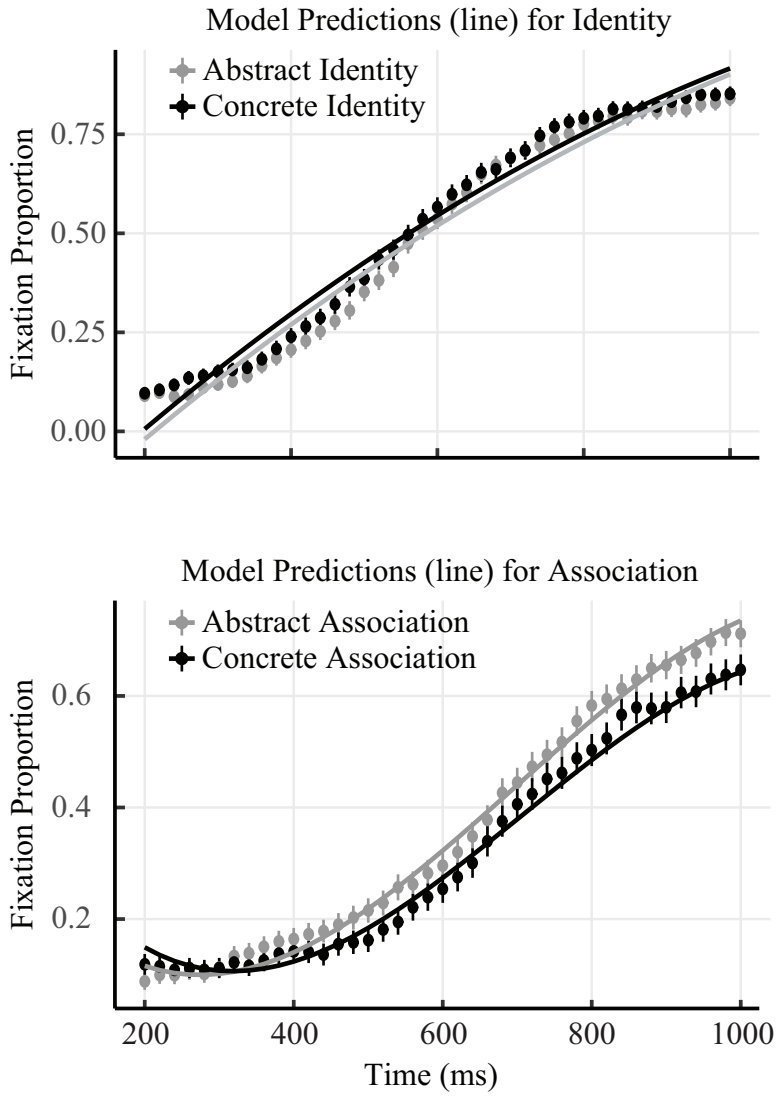

Figure 3. Model predictions for Identity (Top) and Association (Bottom) conditions for 200 ms after auditory word onset. Vertical line throughout each point represents the standard error of the mean

4. Discussion and Conclusion

Investigating the fundamental differences between concrete and abstract concepts is particularly important for understanding the structure of human semantic memory. For a long time, the results of research showed a processing advantage for concrete words (such as 'tree', 'dog', etc.) over abstract words (such as 'victory', 'justice', etc.). Researchers attributed the facilitation effect of concreteness to the quantitative advantages which this type of word possesses (e.g., having direct sensory referents, and the availability of contextual information).

By contrast, more recently it has been proposed that the difference between abstract and concrete words stems from the organizational differences between these types of words. For the first time, Crutch and Warrington (2005) showed that patients with semantic refractory dysphasia demonstrated interference in performing tasks when abstract words were semantically associated rather than semantically synonymous. The pattern for concrete words was reversed, such that concrete words predominantly displayed interference from semantically similar words rather than associated words. This observation, if supported by further research, has important implications for understanding the organizational principle of semantic memory: the differences between abstract and concrete words is qualitative (i.e., concrete concepts are mainly organized based on categorical connections while abstract concepts are organized based on associative connections.).

In the current study, we found that, similarly to Spanish speakers (Duñabeitia et al. 2009), healthy English speakers fixated more on pictures that were associates of abstract words than those which were associates of concrete words. This replication and extension to English supports the existence of qualitative differences between abstract and concrete words. It is worth mentioning that in the current study we considered abstract/concrete dichotomy as a categorical factor, however, concreteness is a continuous variable (Bolongnesi et al. 2020). For the future studies, we need to go further and consider concreteness as a continuous variable to reach a more comprehensive image of organization of these type of words.

The stronger connection found between abstract words and their associates could be related to the way we learn those words. When we learn concrete words, we use actual physical objects in the world as referents. When we learn abstract words, no tangible referents exist, so we more likely learn them through making connections with their associates. We believe that this hypothesis can be tested in the context of second language learning, when the type of word learning can be manipulated by experimenter.

References

Balota, David A., Yap, Melvin J., Hutchison, Keith A., Cortese, Michael J., Kessler, Brett, Loftis, Bjorn, Neely, James H., Nelson, Douglas L., Simpson, Greg B., and Treiman, Rebecca. 2007. The English Lexicon Project. *Behavior Research Methods*, 39(3): 445–459.

Bates, Douglas M., Maechler, Martin, Bolker, Ben, and Walker, Steve. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1): 1–48.

Bolognesi, Marianne., Burgers, Christian., and Caselli, Tommaso. 2020. On abstraction: decoupling conceptual concreteness and categorical specificity. *Cognitive Processin* 21: 365-381.

Bransford, John D., and McCarrell, Nancy S. 1974. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In *Cognition and the symbolic processes,* ed. Walter B. Weimer, and David S. Palermo, 299-303. Hillsdale, NJ: Lawrence Erlbaum.

Cooper, Roger M. 1974. The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology* 6(1): 84-107.

Crutch, Sebastian J. 2006. Qualitatively different semantic representations for abstract and concrete words: Further evidence from the semantic reading errors of deep dyslexic patients. *Neurocase* 12: 91–97.

Crutch, Sebastian J., and Warrington, Elizabeth K. 2005. Abstract and concrete concepts have structurally different representational frameworks. *Brain* 128(3): 615–627.

Duñabeitia, Jon Andoni, Avilés, Alberto, Alfonso, Olivia, Scheepers, Christoph, and Carreiras, Manuel. 2009. Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition* 110(2): 284–292.

Duñabeitia, Jon Andoni, Crepaldi, Davide, Meyer, Antje S., New, Boris, Pliatsikas, Christos, Smolka, Eva, and Brysbaert, Marc. 2018. MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology* 71(4): 808–816. https://doi.org/10.1080/17470218.2017.1310261

Holcomb, Phillip J., Kounios, John, Anderson, Jane E., and West, W. Caroline. 1999. Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25(3): 721–742.

Kieras, David. 1978. Beyond pictures and words: Alternative information processing models for imagery effects in verbal memory. *Psychological Bulletin* 85: 532–554.

Matin, Ethel, Shao, K. C., and Boff, Kenneth R. 1993. Saccadic overhead: Information-processing time with and without saccades. *Perception and Psychophysics* 53 : 372–380.

Mirman, David, Dixon, James A., and Magnuson, James S. 2008. Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language* 59(4): 475–494.

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

Paivio, Allan. 1991. Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology* 45: 255–287.

R Core Team. *R: A Language and Environment for Statistical Computing.* https://www.R-project.org/

Salverda, Anne Pier., Kleinschmidt, Dave, and Tanenhaus, Michael K. 2014. Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of Memory and Language* 71(1): 145–163.

Schwanenflugel, Paula J., and Stowe, Randall W. 1989. Context Availability and the Processing of Abstract and Concrete Words in Sentences. *Reading Research Quarterly* 24(1): 114–126.

Tanenhaus, Michael K., Spivey-Knowlton, Michael J., Eberherd, Kathleen M., and Sedivy, Julie C. 1995. Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science*, 268: 1632–1634.

West, W. Caroline, and 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.