UDK 811.111'373.422
811.111'23
81'373.422
Izvorni znanstveni članak
Prihvaćeno za tisak: 2. lipnja 2025.
https://doi.org/10.22210/suvlin.2025.099.03

Janja Čulig Suknaić, Dora Uštulica
Filozofski fakultet, Sveučilište u Zagrebu
jaculig@m.ffzg.hr, doraustulica@gmail.com

Psycholinguistic variables in lexical–semantic processing of antonyms

In this study we have aimed to examine the role of psycholinguistic variables in the lexical–semantic processing of antonyms. The focus of the study was on the impact of convention on processing antonym pairs, as expressed with the psycholinguistic variable of reaction time. The aim was to further clarify antonymy, not only as a linguistic, but also as a conceptual phenomenon. We used a reaction time experiment to test processing of three different categories of antonyms. Namely, those classified, upon doing a corpus analysis to determine their respective frequencies, as highly conventionalized, less conventionalized and unconventional. Our experiment has confirmed that there is indeed correlation between convention, as expressed by frequency, and reaction time. We take that as just another confirmation of antonymy being not only a lexical but a conceptual phenomenon too, in which both the knowledge of the language and the knowledge of the world play an important part. Finally, we have laid out some of the ways in which this linguistic and conceptual phenomenon could be further studied to get more reliable results to account for the observed effects.

1. Introduction

This study investigates the role of psycholinguistic variables in the lexical–semantic processing of antonyms, with a specific focus on how conventionality influences the processing of antonym pairs. Here we consider the co–occurrence frequency of antonym pairs as a proxy for their conventionality, while recognizing that prototypicality and conventionality are influenced by additional factors (Langacker 1987, Žic Fuchs et al 2013). The central psycholinguistic variable under examination is reaction time, which serves as an indicator of how quickly antonyms are processed. By exploring this relationship, the study aims to provide deeper insights into antonymy as both a linguistic and conceptual phenomenon, as well as to elucidate the role of convention in the processing of lexemes.

It is hypothesized that antonym pairs with a higher degree of conventionalization will be processed more rapidly, reflecting their greater accessibility in the mental lexicon. This would suggest that certain antonym pairs are more prototypical than others, depending on their level of conventionalization, leading to easier lexical access. In this study, we approach the degree of conventionalization through corpus–obtained frequencies of lexemes. While understanding that corpus frequencies are not the same as convention and prototypicality, they are often a good proxy for both, as well as an empirically sound way to obtain information of this kind. Since this study is empirical in nature, it is designed to be replicable and reproducible, in line with the principles of open science. Therefore, corpus–based frequencies of lexemes enable the experiment to be reproducible.

The study begins with a theoretical framework that covers key concepts: antonymy, convention, and reaction time as a psycholinguistic variable. Following this, the central experimental portion of the study is presented. In this experiment, participants are tested on their reaction times when processing pairs of antonyms, which are categorized into three groups: highly conventional, less conventional, and unconventional. The experiment seeks to determine the importance of convention in the lexical–semantic processing of antonyms.

Two hypotheses guide this research:

1. H1: Highly conventionalized pairs, as determined by their co–occurrence frequency in a selected corpus, will be processed more quickly than less conventionalized pairs.
2. H2: Unconventional antonym pairs, which are often context–dependent, will be processed the slowest, although they may still be recognized as opposites.

If differences are observed in the processing speeds across these categories, this would suggest the involvement of distinct cognitive mechanisms, thereby supporting the idea that antonyms can vary in prototypicality depending on their level of conventionalization. Consequently, antonymy should be viewed not merely as a relation between lexemes but as a broader conceptual category of semantic opposition.

The study concludes with a discussion of the findings and proposes directions for future research in this area.

2. Theoretical framework

2.1. Antonymy

Antonyms are a fundamental component of everyday language, a point underscored by Lyons (1977: 271), who noted that “binary opposition is one of the most important principles governing the structure of languages.” Despite their ubiquity, the definition of antonyms in linguistics remains contested. They are most commonly defined as lexemes that are semantically opposite (Murphy et al. 2012). Murphy et al. (2012: 2) define oppositeness in antonymy as "having opposed meanings in a given context," which will be discussed further in the following sections. Additionally, Murphy et al. (2012) distinguish logical oppositeness, which implies that one cannot describe a single entity using both lexemes in an antonym pair, exemplified by the pair *long* and *short*. For instance, if something is described as *short*, it cannot simultaneously be *long*.

However, it is crucial to consider that antonyms are not only oppositional but also share a degree of similarity. This dual characteristic is encapsulated in the principles of maximal similarity and minimal difference (Lyons 1977, Cruse 1986, Murphy 2003), further elaborated by Bianchi and Savardi (2006, 2008a) as the principle of invariance and the degree of adequacy. For example, we do not perceive *pen* and *car* as antonyms because they lack sufficient similarity, unlike the pair *man* and *woman*.

The concept of antonymy has been explored from various perspectives, dating back to Aristotle (Murphy et al. 2013: 6, Čulig Suknaić 2024: 18–23). One prominent perspective is rooted in structuralism, which views antonyms as a type of paradigmatic relation in language. In this view, antonyms are part of a set of potentially substitutable expressions, as they constitute words that can replace each other in context (Lyons 1977: 270). Saussure (1956 [1916]) argued that lexemes derive their meanings from their relationships with other words, a concept that is particularly applicable to antonyms, where each lexeme in a pair is defined by its opposite. Within structuralism, antonyms can be categorized into four groups, as outlined by Cruse (1986): complementaries (e.g., *dead–alive*), contraries (e.g., *fast–slow*), reversives (e.g., *fall–rise*), and converses (e.g., *buy–sell*).

Murphy (2003: 170) introduced the Relation by Contrast–Lexicon Contrast (RC–LC) model, which she defines as "a lexical contrast set that includes only word–concepts that have all the same contextually relevant properties but one." This approach differs from previous ones by considering these relations as metalexical, focusing on "conceptual knowledge about words, rather than the lexical or semantic representation of the words" (Murphy 2003: 171). This perspective emphasizes the importance of factors such as morphology, collocational relations, and phonology, in contrast to the purely semantic dimension of antonymy.

The rise of cognitive linguistics in the 1970s brought about cognitive approaches to antonymy. Cruse and Togia (1996) introduced the concept of the schematic ANTONYM domain, wherein content domains are structured by directionally opposed graded properties. Paradis (1997, 2001) expanded this framework by introducing the concepts of GRADABILITY, OPPOSITENESS, and BOUNDEDNESS. For instance, Paradis (2001: 51) differentiates between scalar and non–scalar adjectives: non–scalar ones are structured along a BOUNDED SCALE (e.g., dead–

alive), whereas scalar adjectives are deemed 'implicit comparatives' and structured along an UNBOUNDED SCALE (e.g., short–long).

The cognitive linguistic approach to antonymy, which is adopted in this study, posits that "antonyms are not just relations between lexemes, but relations between different meaning construals" and that they are "remarkably non–arbitrary" (Croft and Cruse 2004: 40–43, 192). This perspective aligns with our research and hypotheses, suggesting that convention plays a significant role in antonym processing. This underscores the cognitive linguistic view that antonyms reflect relations between meaning construals rather than merely lexical relationships, as our understanding of them is shaped by our interaction with both language and the world in which language is used.

This position is further supported by Čulig Suknaić (2024: 265), who concluded that antonymy functions as a conceptual prototype–based category whose members are added according to conventionalized knowledge of the language and the world. In this view, antonymy is a conceptual category of semantic opposition. Her study also demonstrated that this conceptual approach to antonymy applies across typologically different languages, such as English and Croatian, further validating this perspective.

As illustrated in this section, antonymy has undergone numerous definitions and conceptualizations, highlighting the need for ongoing research to refine our understanding of this complex linguistic phenomenon.

2.2. Convention

The notion of convention is central to linguistics, reflecting "the complexity of language seen not only as a cognitive phenomenon or ability but also as central to human interaction" (Žic Fuchs et al. 2013: 66). Since language operates as both a cognitive and social phenomenon, both aspects must be considered when studying linguistic phenomena, including antonymy. Langacker (1987: 488) defines convention as "the degree to which an expression conforms to the linguistic conventions of a language," meaning that a conventional expression is "widely shared by the members of the relevant speech community." He further asserts that convention is "contextual meaning that is schematized to some degree and established as conventional through repeated occurrence" (Langacker 1987: 157). Thus, the level of conventionalization is reflected in the canonicity of antonyms, demonstrably conventionalized through their co–occurrence in corpora.

In this study, the Corpus of Contemporary American English (COCA) is utilized to gauge the level of conventionalization of antonym pairs – through the proxy of frequency – reflecting their prevalence within the speech community. The frequency of co–occurrence of antonym pairs influences their status and canonicity, as Justeson and Katz (1991: 182) suggest, "adjectives may be more or less antonymous rather than simply antonymous or not antonymous." Čulig Suknaić's (2024: 266) research further supports this idea, concluding that antonymy is structured on the conceptual level as a prototype–based category whose members are included through conventionalized knowledge of the individual pairings and antonymic constructions. More prototypical members tend to be more frequent in the corpora and can be found in a wider set of constructions, which can be used to express more basic opposite meanings.

Aligned with this conclusion, this study tests whether more highly conventionalized (and thus more prototypical) antonym pairs, as indicated by their frequency in the corpus, are processed more quickly. This would suggest that different cognitive representations exist for different antonym pairs in the mental lexicon.

2.3. Reaction time

To demonstrate that some antonyms are more conventionalized and that conventionalization plays a crucial role in their cognitive representation, a reaction time experiment was devised to test the hypotheses. Reaction time (RT) experiments are widely used in psychological and psycholinguistic research, dating back to the pioneering work of F.C. Donders in 1868, who identified three types of RT experiments (Donders 1969 [1868]) which will be outlined below.

According to Decrochers and Thompson (2009), several factors influence word processing, including intrinsic properties (e.g., word length), contextual factors (e.g., word frequency in a given text), and psycholinguistic variables (e.g., imageability, concreteness). This study, as previously discussed, focuses on the notion of convention. As Murphy et al. (2013: 13) argue, psycholinguistic studies suggest that "some antonyms are canonical and so presumably represented as pairs in the mind." Numerous psycholinguistic experiments have supported this claim, including elicitation tests (e.g., Deese 1965, Charles and Miller 1989, Paradis & Willners 2006, Paradis et al. 2009, Čulig Suknaić 2024), identification tests (e.g., Herrmann et al. 1997, Gros et al. 1989, Čulig Suknaić 2024), and semantic priming tests (Becker 1980). These experiments, however, have certain limitations, such as the reliance on the investigator's intuition for stimulus selection and potential confounding factors like word length. In this study, such limitations have been addressed as described in the methodology section.

As stated, an online processing experiment is used to examine how convention influences reaction time, or how quickly participants respond to stimuli in the form of antonym pairs with varying levels of conventionalization. As Jiang (2012: 2) explains:

> "The use of RT data is based on the premise that cognitive processes take time and by observing how long it takes individuals to respond to different stimuli or perform a task in different conditions, we can ask questions about how the mind works and infer about the cognitive processes or mechanisms involved in language processing."

In this study, the variation in the convention variable's effect on reaction time will provide insights into the cognitive processes involved in antonym processing.

There are three types of RT experiments: simple RT tasks, where participants react to stimuli; recognition tasks (go–no–go tasks), where participants respond only to one type of stimulus while ignoring others; and choice tasks, where participants select a response upon stimulus presentation, typically by pressing a predetermined key (Baayen and Milin 2010).

Multiple factors can influence reaction time, including physical and mental condition, age, gender, handedness, and cognitive abilities (Lee and Chabris 2013). These factors can be controlled for in the analysis, as they mainly affect the RT average of individual participants (Wagenmakers et al. 2004). Additionally, factors related to lexemes, such as familiarity, word length, neighborhood density, concreteness, imageability, age of acquisition, spelling–sound regularity, affixation, number of meanings, number of associates, and bigram frequency, should be controlled according to the experiment's specific requirements (Jiang 2012). In this study, particular attention has been given to controlling factors critical to ensuring the reliability of the results.

3. Methodology

3.1. Selection of antonym pairs

The antonym pairs used in this study were selected from established linguistic sources, including Lyons (1977), Cruse (1986), and Jones et al. (2012). These pairs were further refined through a search of the Contemporary Corpus of American English (COCA).[1] Co–occurrence data was retrieved using a symmetric ±5–word search window in COCA's collocates function. No specific antonymy frames were used. A total of 30 antonym pairs were chosen: 15 highly conventionalized pairs and 15 less conventionalized pairs. Additionally, 15 unconventional antonym pairs were identified using ancillary antonymy, a method that generates antonyms through association within a sentence context, as defined by Jones (2004).[2] Frequencies were not provided for these pairs, since they serve to highlight contrastive or context–driven uses of antonymy, where frequency is not a meaningful or applicable measure due to the pairs' inherently atypical status.

1 The Corpus of Contemporary American English (COCA) is a 1–billion–word corpus balanced across multiple genres (spoken, fiction, magazines, newspapers, academic texts, and online media) and spanning texts from 1990 to 2019.

2 *Ancillary antonymy* refers to a phenomenon in which an antonymous pair appears in a sentence not as the primary focus of contrast, but as a secondary or supporting element that helps establish or highlight a more significant, often context–specific, opposition between another pair of words, phrases, or clauses. This is illustrated in the following examples from Jones (2002):

(1) I love *to cook* but I hate *doing the dishes* – so I'd have a dishwasher or a family of gypsies to do the washing up.

(2) Since then, of course, they've all had knighthoods, usually when they're too old *to play Hamlet* but too young *to play butlers in Hollywood movies*.

The criteria for determining the level of conventionalization were based on the frequency of antonym pair occurrences in COCA. Highly conventionalized pairs had a frequency of over 2,000 tokens, while less conventionalized pairs had a lower frequency. The threshold of 2,000 tokens was established based on the observed median and average frequency values for antonym pairs in the dataset. While not drawn from a specific prior study, this cutoff point reflects a natural break in the frequency distribution and it served as a practical boundary for distinguishing commonly co–occurring pairs from those with notably lower frequencies. All selected pairs consisted of monosyllabic words to control for the potential confounding effect of word length.

The table below contains all the conventionalized pairs together with their frequencies and also all the chosen unconventional pairs.

HIGHLY CONVENTIONALIZED		LESS CONVENTIONALIZED		UNCONVENTIONAL
black–white	34707	near–far	1017	sun–moon
good–bad	16337	wet–dry	968	book–film
left–right	16212	hard–soft	921	car–bus
come–go	15629	fast–slow	873	sea–land
day–night	10016	laugh–cry	786	faith–doubt
life–death	7799	thick–thin	652	food–drink
high–low	7329	push–pull	578	debt–cash
east–west	5684	sit–stand	520	dress–pants
front–back	4182	tall–short	486	walk–run
long–short	4159	thin–fat	343	greed–need
big–small	3844	well–ill	219	red–blue
boy–girl	3715	pass–fail	199	spoon–fork
poor–rich	3690	lean–fat	181	sun–rain
give–take	3306	dim–bright	74	heart–brain
love–hate	2827	break–fix	34	ice–sand

3.2. Experiment design

The experiment was designed and implemented using PsyToolkit (Stoet 2010, 2017), a free platform for running cognitive–psychological experiments. The full experiment and survey materials are available in the Open Science Framework repository[3].

3 https://osf.io/wqhtd/files/osfstorage

3.2.1. Reaction time experiment

The primary task in the experiment involved measuring participants' reaction times in identifying antonym pairs. Participants were presented with word pairs and instructed to decide whether the pair represented opposites by pressing a designated key. Fifteen (15) synonym pairs were included as a control group. Reaction time was recorded from the moment the word pair appeared on the screen until the participant pressed a key. The stimuli were presented in a fully randomized order for each participant to minimize potential order effects and ensure that no systematic bias influenced response times or accuracy.

To minimize confounding variables, all instructions and stimuli were designed using Inkscape (Harrington et al., 2004–2005). The task falls under the decision RT category, combining elements of both simple and choice reaction time tasks. Each trial began with a fixation point displayed for 500 ms, followed by the presentation of a word pair. Participants had two seconds to respond. The right arrow key was used to indicate that the pair were opposites, and the left arrow key indicated they were not. Feedback was provided after each response, indicating whether the participant was correct, incorrect, or too slow. It was part of the experimental software's default configuration and was not intended to influence participant strategy. We acknowledge that immediate feedback may lead to adaptive behavior during the task and that this could potentially affect response patterns. However, since lexical decisions taken here were binary in design, the feedback served primarily to keep participants engaged and alert throughout the task. While this approach may appear to constrain a more nuanced gradable view of antonymy, it reflects a common experimental simplification used in psycholinguistic studies to enable systematic data analysis (e.g. Cruse 1986; Justeson et al. 1991; Murphy 2003). Future studies may consider omitting trial–level feedback or varying its presence to assess potential adaptive effects more systematically.

The entire experiment lasted approximately three minutes per participant.

3.2.2. Data collection and analysis

The anonymized results for each participant were saved in Excel sheets for subsequent analysis. Incorrect responses were excluded from the analysis. Additionally, reaction times that fell outside the range of 200 to 2000 ms were excluded as outliers, as these were likely due to accidental key presses or participant distraction (Jiang, 2012).

3.3. Participants

The study included 50 participants, with a mean age of 38 (range: 19–73). The gender distribution was 32 females and 18 males. Educational backgrounds varied, with 22 participants having a high school diploma, 20 holding a bachelor's degree, 6 having a master's degree, and 2 holding a PhD. All participants were native English speakers, with 46 participants from the UK and 4 from the USA. Only one participant reported having a language disorder, two were unsure, and the remaining participants reported no language disorders.

3.3.1. Participant demographics

Participants were first asked to complete a brief survey collecting basic demographic information, including gender, age, education level, possible language disorders, native language, and country of origin. This initial survey also included a brief description of the experiment, an informed consent form, contact information for the researcher, a data anonymization statement, and detailed instructions for the upcoming experiment.

4. Results

There is still no consensus on the best way to analyze reaction time data, as they have their specific features, which will be discussed in this section, and consequently different researchers choose to analyze reaction time data differently. So, before starting the analysis, some of the methodological concerns regarding the analyses of RT data will be presented.

The first problematic point is that distributions of RTs are often positively skewed, resulting in non–normal distributions (Baayen and Milin 2010). This problem may be solved by conducting non–linear transformations of data, transforming the data logarithmically. However, recently it has been shown that this may not be really useful if one only wants to see whether RT covariates with variables such as frequency (Schramm and Rouder 2019). The second problem is that individual responses are not statistically independent, as a trial–by–trial sequential correlation is present (Baayen and Milin 2010). That is why it is best to account for the individual differences in the model as well.

When it comes to the distribution of the data, neither the mean, nor the standard deviation are robust measures because the distribution is often skewed. That is why many researchers choose to report the median value as a central tendency parameter and interquartile range for estimating dispersion. This is also not a perfect way to analyze data, since the median is a biased estimator when the data is skewed, resulting in underestimation (Whelan 2008). One of the proposed solutions to this is to analyze central tendencies of the whole distribution to better understand the differences between different categories (Whelan 2008).

As can be seen, there are many possible obstacles one can stumble upon when analyzing RT data. However, led by the position articulated by the statistician George Box (1976), which could be paraphrased as saying that all statistical models are wrong, but some are useful, we have analyzed the results of the experiment with the intention of getting the most out of them, while always being aware of their possible limitations.

All the results were converted into .csv format. They were imported into RStudio (version 1.4.17) and analyzed with R (version 4.1.0). The analysis was done primarily with the tidyverse (Wickham et al. 2019), patchwork: The Composer of Plots (Lin Pedersen 2020), mgcv (Wood 2011) and lme4 (Bates et al. 2015) packages. The code used for the analysis can be accessed in either HTML or R Markdown form via the Open Science Framework link provided.

4.1. Descriptive statistics

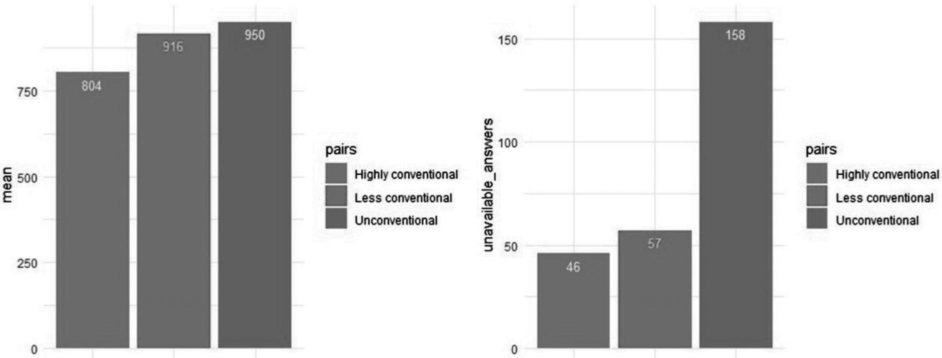

Figure 1: RT mean (left) and number of NAs (right)

In the figure above two bar plots are presented. The first one represents the distribution of mean RTs per different experimental categories. Namely, highly conventional pairs, less conventional pairs and unconventional pairs. All the values were rounded. Highly conventional pairs were on average processed the fastest by the participants (M = 804, SD = 240, Mdn = 757, range = 324 – 1949, IQR = 256), followed by less conventional pairs (M = 916, SD = 280, Mdn = 878, range = 273 – 1967, IQR = 347). Finally, unconventional pairs took participants the longest to process (M = 950, SD = 287, Mdn = 923, range = 271 – 1845, IQR = 371). What can be observed right away is that the means in all cases are greater than the medians, and that there are data points that exceed two standard deviations, which tells us that the data is right skewed. We can also see that interquartile ranges are ranked in the same way as means, from the lowest for the highly conventionalized pairs to the highest for the unconventional pairs. That tells us that the dispersion in the middle half of our distribution, which is especially important in skewed distributions, is the greater the less conventional the pairs are. We have conducted two–sample T–tests to determine whether the differences between the means are in this case statistically significant. The mean difference between highly conventional and less conventionalized pairs turned out to be statistically significant ($p = 0.03$, $d = 0.43$, $t = -2.15$), just like the difference between highly conventional and unconventional pairs means ($p = 0.0069$, $d = 0.55$, $t = -2.76$). The difference between the less conventionalized pairs and unconventional pairs is not statistically significant and the magnitude of difference is small ($p = 0.55$, $d = 0.12$, $t = -0.59$). However, that does not mean that the difference between these two groups is insignificant. That is why we also observed the unavailable answers for each category. The term "unavailable answers" was adopted as it aligns with the default categorization used in the data retrieval software, streamlining data processing and ensuring consistency. Although incorrect responses and those outside the reaction time scope represent conceptually different categories, they share a crucial commonality – both types of responses produce data that are unsuitable for inclusion in the primary analysis of valid and timely answers. Incorrect responses were not recognized by the participants as antonym pairs, while responses outside the RT limits violated temporal constraints required for meaningful interpretation. Grouping them together under "unavailable answers" reflects their shared role as non–interpretable or invalid data points within the subsequent inferential statistics analysis, facilitating a clearer data cleaning process and allowing for focused interpretation of valid responses.

In the bar plot on the right–hand side of the graphic, we can see the distribution of unavailable answers per each category. Those were the answers that either took the participants too long to answer or they did not give the correct answer. In the case of our study, it would be wrong to simply ignore those answers because they are informative. We can see right away that the difference between them is significant. 17.62% of those answers can be found in the group of highly conventional pairs, 21.84% in the group of less conventional pairs, and 60.54% in the group of unconventional pairs. Two sample T–tests were also conducted on them. The difference between the highly conventionalized and less conventionalized pairs was not significant ($p = 0.12$, $d = 0.32$, $t = -1.58$). However, tests for highly conventionalized and less conventionalized pairs paired with the unconventional pairs were statistically significant, with large magnitude of difference effects ($p < 0.001$, $d = 1.57$, $t = -7.84$; $p < 0.001$, $d = 1.57$, $t = -7.83$).

What we can conclude upon analyzing the distributions of data is that, when it comes to reaction time, highly conventionalized pairs stand out, as they are supposedly easier to access and subsequently to process. The difference between the less conventional and unconventional pairs is not significant, which suggests that, when unconventional pairs are recognized as antonyms, they are equally easy, or rather, equally difficult to access as the less conventional pairs. This underscores the fact that, when it comes to lexical access, the degree of conventionalization should be high for it to matter. If we observe the number of unavailable answers, the conclusion is opposite. Namely, there is no significant difference between the highly and less conventionalized pairs, which means that the latter are recognized as antonyms, just not as easily. Unconventional pairs are sometimes not recognized as antonyms, or would take much longer to process, specifically because they are not conventional. Another observation to be made is that unconventional antonyms are still very often recognized as antonyms by the majority of participants (in our sample they appeared 750 times, but there were only 158 unavailable answers for them). This tells us that antonyms, as we defined them here, should primarily be regarded as a relation of conceptual opposition, and not only as lexical opposition.

After conducting the initial analysis on the whole sample and per category, we have decided on the further path of analysis. Namely, outlier values exceeding 2 SDs will not be removed, as in the case of our research they affect both the results, and the assumptions made. Unavailable answers are going to be excluded from the analysis after we proved that the difference in means is significant even without them being replaced by means per category or by regression imputation. Instead, the number of unavailable answers is going to be included in the mixed–effects model to account for that factor. Finally, we have decided not to conduct a non–linear transformation on our data, as it would not significantly affect the analysis.

4.2. Plots per item

To get a better picture of our data, we have made kernel density graphs for every category of antonym pairs. They show the distribution of every item across participants. A vertical line was added to each of them, to mark the mean of the distribution.

4.2.1. Highly conventionalized pairs

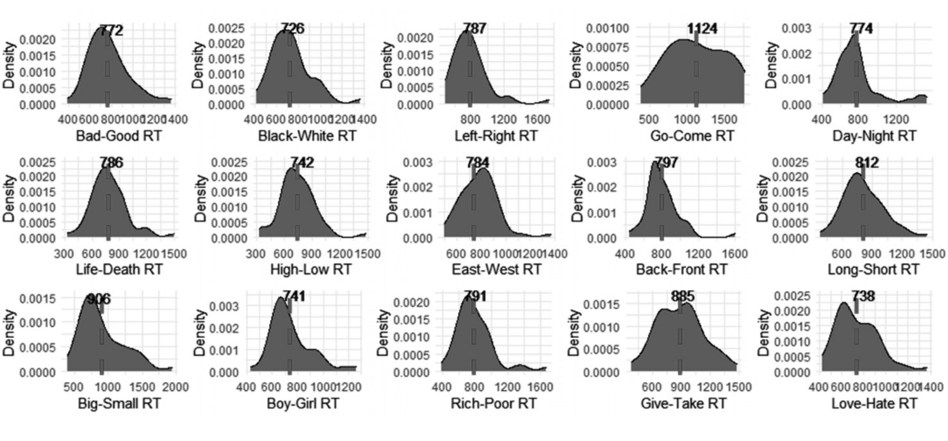

Figure 2: Kernel density graphs – highly conventional pairs

If we look at the graphs above, we can see that almost all of them are positively skewed, which was expected, since the means tend to be greater than the medians in this type of data because of the outliers. Some cases are interesting to observe. Firstly, the pair which has the highest RT, *go–come,* is dispersed across the RT distribution, which means that for some participants it was rather easy to process, and for others it was rather difficult. Bimodal distributions can be seen in pairs *love–hate, give–take, boy–girl*, and *big–small*. So, there was a significant number of participants who had longer RTs than the observed mean, or rather, the distribution of RTs is dispersed with significant individual differences. The examples which have a distribution similar to the normal distribution are *bad–good, high–low,* and *long–short*. They were on average similarly difficult to process for all participants.

4.2.2. Less conventional pairs

The graphs below represent less conventional pairs. The pairs that have a distribution similar to the normal one are *cry–laugh, lean–fat, fix–break,* and *ill–well*. They are the pairs that on average had longer processing times. So, we can say that they were on average equally difficult to process for the participants. The pair that took the longest to process, *dim–bright*, does have a distribution similar to the normal one, but there was a smaller number of participants for whom it was even harder to process, so it can be considered bimodal. If we look at the other pairs with bimodal distributions, this type of distribution can be seen in the examples *near–far* and *tall–short*. These examples were then also considerably harder to process for a number of participants.

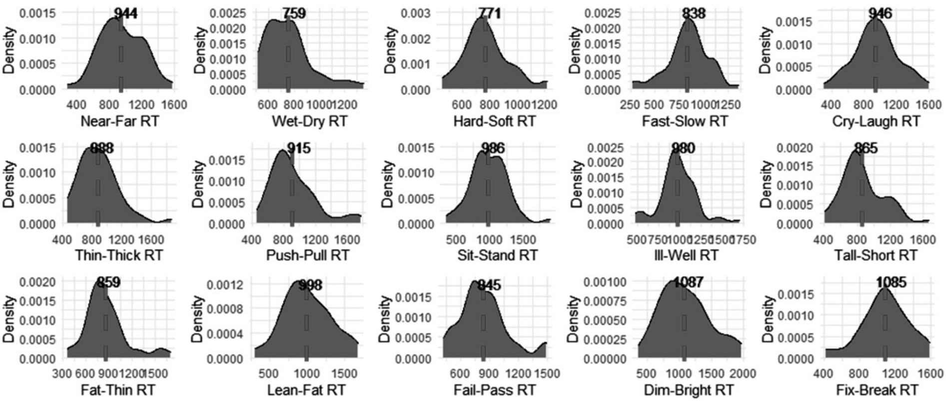

Figure 3: Less conventional pairs – kernel density graphs

4.2.3. Unconventional pairs

In the case of unconventional pairs, we can see that the majority of graphs are actually similar to normal distributions. So, in their case, there were not that many individual processing differences among the participants. There are no graphs with bimodal distributions, and the means and medians of this group of pairs are close. This is why there are not many examples that are significantly right skewed. The only pairs that have a somewhat more distributed dispersion are *car–bus, need–greed*, and *dress–pants*.

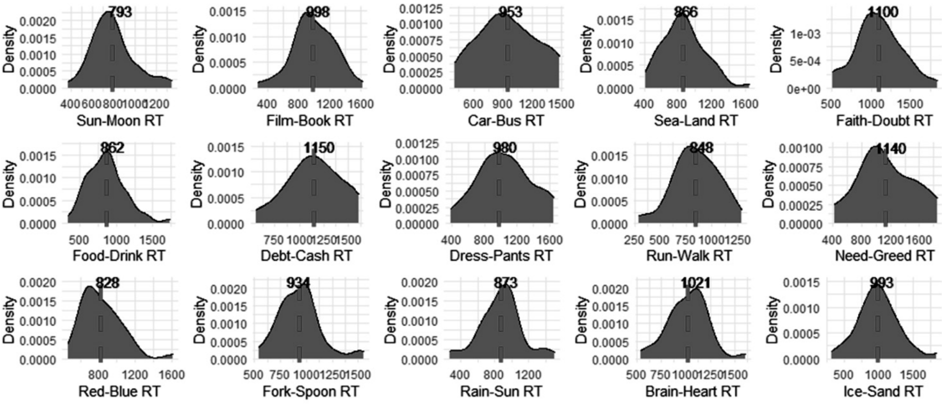

Figure 4: Unconventional pairs – kernal density graphs

Upon observing these kernel density graphs, we can conclude that data like this is dominated by individual differences, which should ideally be taken into account in statistical modeling. We can also conclude that, although the outliers that skew the graphs are present in the majority of them, there is still a number of normal distribution graphs that can be observed, as well as a number of bimodal ones, so it was not necessary to log–transform the data. The outliers can in this case also be informative. For instance, we have observed that they were less present in the group of unconventional pairs than in the other two, which tells us that their processing is more averaged. On the other hand, they were prevalent in the group of highly conventionalized pairs.

4.3. Inferential statistics

To generalize our conclusions to the population level we have fitted several generalized additive and mixed–effects models. In generalized additive models we have looked at the correlation between frequencies and RTs. RT is modelled on the y–axis as the dependent variable and frequency is modelled on the x–axis as the predictor variable. In mixed–effects models we have added the number of unavailable answers as a random variable to the model.

To justify the use of errors as a random variable, we have fitted a linear regression model to show the correlation between the number of unavailable answers and RT. As can be seen in the graph below, there is a clear positive correlation between the two. That was confirmed in the summary of the model ($p < 0.01$, adjusted R–squared = 0.51, Pearson's $r = 0.71$). This tells us that those pairs that had more unavailable answers were also the ones that took participants longer to process, which means that accuracy and lexical access are also connected. This also means that, if we were to transform the data of the unavailable answers, they would have to be transformed carefully, as those answers would on average take longer than the mean of the group, which is often used to replace them.

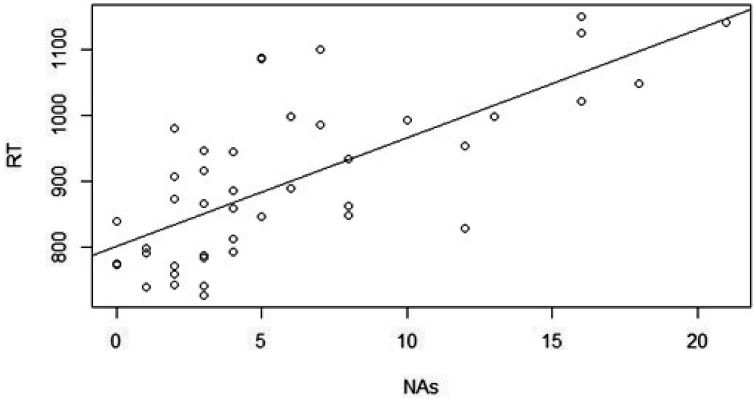

Figure 5: Linear regression model of RT and NAs

4.3.1. *Highly conventional pairs*

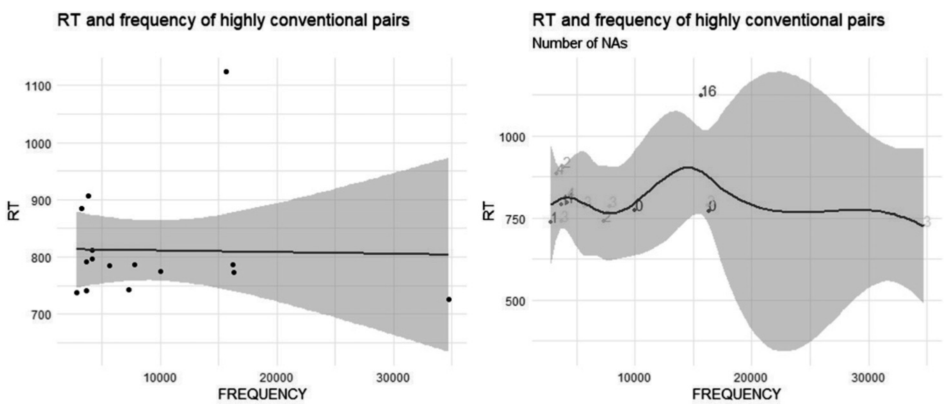

Figure 6: GAM of highly conventional pairs (left) and mixed–effects model of highly conventional pairs (right)

The first model that can be seen above on the left–hand side is the generalized additive model (GAM) of the highly conventional group of pairs. As can be seen, the regression line is flat and we can immediately conclude that the correlation between frequency and RTs in this case is not significant, which was also confirmed by performing the summary of the model ($p = 0.93$, adjusted R–squared = –0.076, deviance explained = 0.686%, Pearson's $r = -0.025$). To the right of the GAM we can see a mixed–effects model, in which we included the errors made per every pair. We can see that the only outlier with the highest RT is the one which has 16 unavailable answers and that is the pair *come–go*. To sum up, frequency is not in correlation with RT in the highly frequent data. This also confirms that the cutoff value of 2000 tokens chosen for the highly conventional pairs was a reliable one, as participants did not seem to perceive the differences between the pairs within this group.

4.3.2. Less conventional pairs

The first model that can be seen below on the left–hand side is the GAM of the less conventional group of pairs. As can be seen, the regression line in this case is not flat, so we can immediately conclude that correlation between frequency and RT is in this case present. The summary of the model has confirmed that there is a strong to moderate negative correlation between the two (p = 0.007, adjusted R–squared = 0.39, deviance explained = 43%, Pearson's r = – 0.66). So, in the case of less conventional pairs, frequency is in a strong negative correlation with RT. To the right of the GAM we can see a mixed–effects model, with the unavailable answers made per every pair just like in the previous example. In this case there is much more variation. There are no particular outliers, meaning that there were no great differences in accuracy in this group.

We can conclude that, when it comes to the pairs that do not have a high, or rather a high enough frequency, differences between different pairs do matter. The range between the frequencies of the pairs of highly conventional antonyms is greater (2,827–34,307; the difference being 31,480 tokens) than that of the less conventional pairs (34–1,017; the difference being 983 tokens), but no correlation was found. It seems that, after a certain level, the difference does not matter, as we simply do not perceive it.

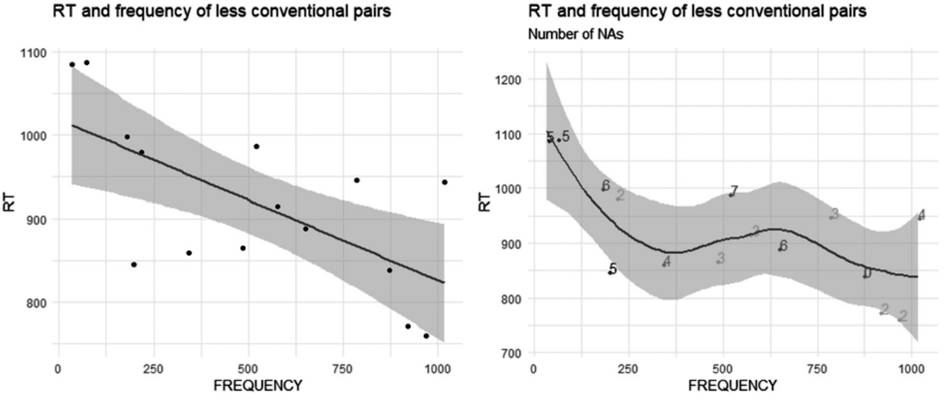

Figure 7: GAM of less conventional pairs (left) and mixed–effects model of less conventional pairs (right)

4.3.3. All pairs combined

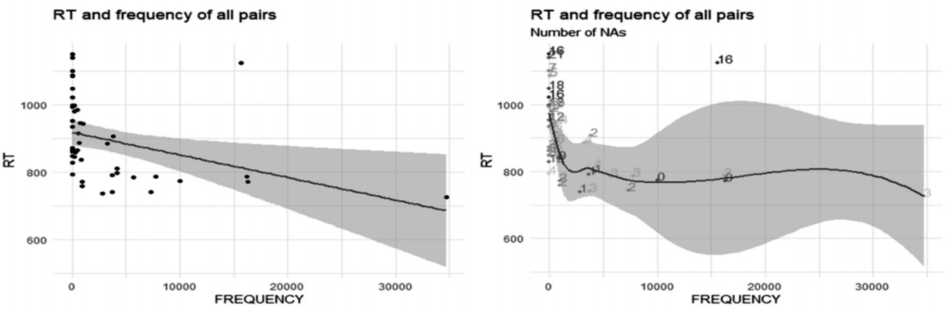

Figure 8: GAM of all pairs (left) and mixed–effects model of all pairs (right)

The first model that can be seen above on the left–hand side is the GAM of all pairs combined. The frequencies in the case of the unconventional pairs were set to zero, and that is why data points are concentrated around it. As can be seen, the regression line in this case is also not flat, so we can conclude that, once again, this is a case of negative correlation. The summary of the model has confirmed that (p = 0.01, adjusted R–squared = 0.11, deviance explained = 12.8%, Pearson's r = – 0.36). In this case, correlation is not as strong as in the previous case. It is a low to moderate one. Still, we can say that frequency, or rather convention as we have modelled it on the basis of corpus–obtained frequencies of lexemes, is indeed associated with reaction time. If we look at the mixed–effects model, we can see that the data points per category with the highest number of unavailable answers are indeed placed on the highest points of the y–axis, indicating high RTs. Those are the following pairs: *need–greed* (21), *dress–pants* (18), *brain–heart* (16), *debt–cash* (16), *come–go (16), faith–doubt* (7) and *dim–bright* (5).

We can say that the degree of conventionalization, which we have modelled as frequencies derived from the corpus analysis, does in fact influence reaction time (RT). This can be confirmed both by looking at the models of the highly conventional pairs, in which no correlation between the two was found, and which only underscores the fact that those pairs are really highly conventionalized (or rather, prototypical), and at the model of less conventional and all pairs combined where we have observed that frequency does indeed influence lexical–semantic processing. However, there are multiple other factors that should be considered, and they are going to be discussed in the following section.

5. Discussion

Our study provides empirical support for both proposed hypotheses. We found that antonym pairs classified as highly conventional are processed significantly faster than those classified as less conventional. Conversely, unconventional antonym pairs are processed the slowest. This pattern was statistically significant, as confirmed by the combination of p–values and Pearson's *r*. However, the R–squared adjusted values were not very high, indicating that while our model accounts for some variance, it may not capture all aspects of antonym processing. This aligns with the observation that "language is complex and humans are messy" (Winter 2020: 77), suggesting that linguistic models often fall short of explaining every nuance.

5.1. Methodological considerations

Data Analysis and Model Limitations: Our chosen method of analysis, while informative, may not fully capture the high level of micro–variation present in both participant responses and individual items. A multilevel analysis, such as a Bayesian approach, could better account for these variations and might be a more suitable analytical tool for future studies.

Factors Influencing Reaction Time: Several factors could affect reaction time, including participant age and individual differences. Given the broad age range of our participants (19–73 years), incorporating age as a random variable in the model might yield more nuanced insights. Additionally, while we collected data on potential language disorders, the low incidence suggests that including this factor might not have had a significant impact. However, in future research, it would be prudent to account for such variables comprehensively.

Word Frequency and Related Features: Word frequency, used as a proxy to conventionalization, is associated with various other lexical features, such as word length, age of acquisition, and word similarity. Brysbaert et al. (2018) highlight that these factors can act as confounding variables. For example, our study observed that low–frequency antonyms sometimes had processing times comparable to, or faster than, high–frequency antonyms (e.g. *hard–soft* vs. *high–low*). This suggests that while frequency is a significant factor, it does not always predict processing speed. Brysbaert et al. (2016) found that word frequency accounts for only 30–40% of the variance in lexical processing, indicating that other factors also play a crucial role. To address this, future studies could incorporate additional variables, such as word prevalence and participant familiarity with specific antonym pairs. Implementing a pre–experiment questionnaire to assess participants' knowledge of the antonym pairs could provide valuable data triangulation.

Psycholinguistic Variables: Besides frequency, other psycholinguistic variables—such as concreteness, imageability, and valence—could influence reaction times. These variables were not directly measured in our study, but could be explored in future research to gain a more comprehensive understanding of antonym processing.

5.2. Future directions

Alternative Methods: Reaction times, while useful, are indirect indicators of cognitive processes. To obtain a deeper understanding of antonym processing, employing additional methods, such as EEGs or fMRIs, could offer more direct insights into brain activity and cognitive mechanisms. These methods have their own specificities, but could complement reaction time studies to further investigate the relationship between antonymy and convention.

In summary, our study supports the hypothesis that antonym pairs with higher conventionalization are processed faster, while unconventional pairs are processed more slowly. Although our findings are significant, they also highlight the complexity of linguistic processing and the need for more refined analytical methods and additional psycholinguistic measures. Future research should consider these factors to build a more comprehensive model of antonym processing and its underlying cognitive mechanisms.

6. Conclusion

This study explored the relationship between conventionalization—measured by the frequency of antonym pairs—and their lexical–semantic processing. Our findings support both of our hypotheses: antonyms classified as highly conventional were processed significantly faster, while unconventional antonyms were processed more slowly. This was also reflected in the number of unavailable answers for each category, suggesting that the level of conventionalization influences how readily antonyms are accessed in the mental lexicon.

Our results further affirm the cognitive linguistic perspective that antonyms are not only lexical but also conceptual opposites. Participants were able to recognize even unconventional antonyms as opposites, indicating that antonym processing involves both linguistic knowledge and conceptual understanding. This aligns with the view that antonyms are represented in our mental lexicon based on their conventional usage and conceptual relationships.

Given the complexity of the interaction between antonym conventionalization and processing, future research should consider additional factors discussed in this study. For instance, incorporating variables such as participant demographics and psycholinguistic features could provide deeper insights into these dynamics.

In summary, our study highlights the interplay between convention and lexical–semantic processing of antonyms, reinforcing the notion that language processing is a multifaceted cognitive and social phenomenon. Continued investigation into this area will contribute to a more nuanced understanding of how antonyms are conceptualized and processed in the human mind.

References

Baayen, R.H. and Milin P. (2010). "Analyzing Reaction Times". In: *International Journal of Psychological Research, 3 (2), 12–28*.

Bates D., Machler M., Bolker B., Walker S. (2015). "Fitting Linear Mixed–Effects Models Using lme4". In: *Journal of Statstical Software*, 67 (1), 1–48.

Becker, Curtis A. (1980). "Semantic context effects in visual word recognition: an analysis of semantic strategies". In: *Memory and Cognition* 8, 493–512.

Bianchi I. & Savardi U. (2006). "Oppositeness in visually perceived forms". In: *Gestalt Theory*, 4, 354–374.

Bianchi I. & Savardi U. (2008). *The perception of Contraries*. Aracne, Roma.

Box, George E.P. (1976). "Science and Statistics". In: *Journal of the American Statistical Association*, 71(356), 791–799.

Brysbaert, M., & Ellis, A. W. (2016). "Aphasia and age of acquisition: Are early–learned words more resilient?" In: *Aphasiology* 30, 1240–1263.

Brysbaert M, Mandera P, Keuleers E. (2018). "The Word Frequency Effect in Word Processing: An Updated Review". In: *Current Directions in Psychological Science* 27(1). 45–50. doi:10.1177/0963721417727521

Croft, William and D.Alan Cruse. (2004). *Cognitive linguistics.* Cambridge University Press.

Cruse, D.Alan. (1986). *Lexical semantics*. Cambridge University Press.

Cruse, D.Alan and Pagona Togia. (1996). "Towards a cognitive model of antonymy". In: *Journal of Lexicology* I, 113–41.

Čulig Suknaić, Janja. (2024). *Značenjska suprosnost. O antonimiji u engleskome i hrvatskome jeziku*. Hrvatska sveučilišna naklada.

Deese, James.(1964). "The associative structure of some common English adjectives". In: *Journal of Verbal Learning and Verbal Behavior* 3, 347–57.

Descrochers, Allan and Glenn L. Thompson. (2009). "Subjective frequency and imageability ratings for 3,600 nouns". In: *Behavior Research Methods* 41 (2): 546–557. https://doi.org/10.3758/BRM.41.2.546.

Donders, Franciscus. (1969). "On the speed of mental processes". In: *Acta Psychologica*, 30 , 412–431. (Translated by W. G. Koster)

Harrington, B. et al. (2004–2005). *Inkscape*.

Jiang, Nan. (2012). *Conducting Reaction Time Research in Second Language Studies*. Routledge.

Jones, Steven. (2004). *Antonymy. A corpus–based perspective.* Taylor & Francis e–Library.

Jones, Steven, M. Lynne Murphy, Carita Paradis, and Caroline Wilners. (2012). *Antonyms in English:construals, constructions and canonicity.* Cambridge University Press.

Justeson, John S. and Slava M. Katz. (1991). "Co–occurrences of Antonymous Adjectives and Their Contexts". In: *Computational Linguistics*, 17 (1):1–20.

Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar*. Stanford University Press

Lee, J.J. and C.F.Chabris. (2013). "General cognitive ability and the psychological refractory period: Individual differences in the minds bottleneck". In: *Psychological Science* 24 (7): 1226–1233.

Lehrer, Adrienne and Keith Lehrer. (1982). "Antonymy". In: *Linguistics and Philosophy* 5, 483–501.

Lyons, John. (1977). *Semantics.* Cambridge University Press.

Murphy, M.Lynne. (2003). *Semantic relations and the lexicon.* Cambridge University Press.

Paradis, Carita. (1997). *Degree modifiers of adjectives in spoken British English.* Lund University Press.

Paradis, Carita. (2001). "Adjectives and boundedness". In: *Cognitive Linguistics* 12, 47–66.

Paradis, Carita and Caroline Willners. (2006). "Antonyms and negation: the boundedness hypothesis". In: *Journal of Pragmatics* 38, 1051–80.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R–project.org/.

RStudio Team. (2021). *RStudio: Integrated Development Environment for R. RStudio*, PBC, Boston, MA URL: http://www.rstudio.com/.

Saussure, Ferdinand de. (1956 [1916]). *Course in general linguistics.* London: Peter Owen.

Schramm, Pele & Rouder, Jeffrey. (2019). *Are Reaction Time Transformations Really Beneficial?*.10.31234/osf.io/9ksa6.

Stoet, Gijsbert. (2010). "PsyToolkit – A software package for programming psychological experiments using Linux". In: *Behavior Research Methods, 42(4)*, 1096–1104.

Stoet, Gijsbert (2017). "PsyToolkit: A novel web–based method for running online questionnaires and reaction–time experiments". In: *Teaching of Psychology*, 44(1), 24–31.

Wagenmakers E., S. Farrell, and R.Ratcliff. (2004). "Estimation and interpretation of 1/fα noise in human cognition". In: *Psychonomic Bulletin & Review* 11, 579–615.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. (2019). "Welcome to the tidyverse". In: *Journal of Open–Source Software*, 4(43), 1686. doi:10.21105/joss.01686.

Winter, Bodo. (2020). *Statistics for Linguists: An Introduction Using R*. Routledge: New York.

Whelan, Robert. (2008). "Effective analysis of reaction time data". In: *The Psychological Record* 58, 475–482.

Wood SN. (2011). "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models". In: *Journal of the Royal Statistical Society* (B), 73(1), 3–36.

Žic Fuchs, Milena, Vlatko Broz and Nina Tuđman Vuković. (2013). "Communication technologies and their influence on language: the notion of convention revisited". In: *Jezikoslovlje* 14, 65– 84.

Corpus

COCA (Corpus of Contemporary American English) https://corpus.byu.edu/COCA/

Psiholingvističke varijable u leksičko–semantičkoj obradi antonima

Cilj je ovog istraživanja propitati ulogu psiholingvističkih varijabli u leksičko–semantičkoj obradi antonima. Istraživanje je usmjereno na utjecaj konvencije i razine konvencionalizacije na obradu parova leksema suprotnoga značenja uz pomoć psiholingvističke varijable vremena reakcije. S tako usmjerenim istraživanjem želi se pobliže rasvijetliti semantički karakter antonimije, ne samo kao jezične nego i kao pojmovne pojavnosti. Stoga je za ostvarivanje cilja odabran eksperiment mjerenja vremena reakcije, kojim se ispituje obrada triju različitih vrsta značenjskih odnosa suprotnosti, to jest antonimije, a koji omogućuje uvid u obradu: visoko konvencionalizirani antonimi, nisko konvencionalizirani antonimi i nekonvencionalizirani antonimi. Za potrebe eksperimenta najprije je provedeno korpusno istraživanje, kojim su se utvrdile frekvencije odabranih parova antonima, što je poslužilo za određivanje razine njihove konvencionaliziranosti. Visoko frekventni parovi određeni su kao visoko konvencionalizirani, a nisko frekventni kao nisko konvencionalizirani, i to zbog mjerljivosti i potrebe za kvantifikacijom rezultata uslijed postavljanja eksperimenta. Provedeni eksperiment vremena reakcije pokazao je da postoji korelacija između konvencije (određene na temelju frekvencije unutar korpusa) i vremena reakcije. Rezultat eksperimenta smatramo još jednom potvrdom tvrdnje da antonimija nije jednostavna leksička kategorija, nego složena pojmovna pojavnost pri čijoj obradi govornik ovisi o svojem složenom i isprepletenom znanju o jeziku i znanju o svijetu. Ovo istraživanje ukazuje na potrebu za daljnjim koracima u svrhu produbljivanja znanja o antonimiji kao jezičnoj i pojmovnoj pojavnosti, a koji bi mogli pomoći u interdisciplinarnim uvidima u procese ljudskoga mišljenja, gradbe znanja i obrade jezika.

Ključne riječi: antonimija, konvencije (semantika), vrijeme reakcije, psiholingvističko istraživanje, engleski jezik

Key words: antonymy, conventions (semantics), reaction time, psycholinguistic research, English