What Word-Guessing Reveals About Your Brain:

Patterns of Lexical Storage and Processing

Kamila Kuishibekova

Department of Languages, Linguistics and Literature, Nazarbayev University

WLL 499: Languages, Linguistics, & Literatures Capstone

Instructor: Professor Eva-Marie Dubuisson

Capstone Advisor: Professor Valentina Apressyan

May 4, 2025

Table of Contents

Introduction

There are three theories in psycholinguistic research that concern the structure of the mental lexicon, which is the mental representation of words and meanings in the human mind. These are the language-distributional models, that suggest defining the word meanings by their co-occurrences with other words, attributional theories, where concepts that share many similar features are stored closer, and the hybrid theories. The hybrid theories offer an inclusive approach, suggesting it is a combination of attributional and distributional information that develops our rich semantic system (Andrews, Frank, & Vigliocco, 2014, as cited in Meteyard & Vigliocco, 2018).

By analyzing how players of Alias game use language strategies, this study aims to make three key contributions: (1) introduce a naturalistic gameplay method, (2) provide evidence for hybrid semantic organization, and (3) explore how word-level features modulate lexical storage and processing.

Alias is a word-guessing game, where players team into pairs of a speaker and listener, and guess words. The speaker describes as many words (out of 8 in the Russian edition of the game) on the game card as they can, and the listener makes guesses based on the speaker's clues. Playing Alias, people use various creative strategies to generate clues, such as formulating definitions ('a group of people traveling together, often in a desert'), providing synonyms (cavalcade, expedition, procession), providing associations (desert, camels, sands) or providing the initial part of a frequent fixed word combination (collocation) in order to elicit the target word (*The dogs bark but … goes on*).

Word association games like Alias involve both active language comprehension and language production under time constraints. Clark (1971) identifies three stages in such games: (1) the player must understand the stimulus, (2) process its meaning, and (3) produce a response. In addition to these cognitive processes, Alias requires players to optimize their chances of success by considering shared knowledge and other strategic factors. This dynamic setting makes Alias a valuable tool for studying theories of retrieval and production. By employing a game-based research design, this study analyzes natural speech in a setting with high ecological validity.

Literature Review

Introduction to the Mental Lexicon

The mental lexicon is not just a dictionary-like storage of word meanings. We store not only the semantic information about the words: we know how to pronounce them, we recognize them visually and have awareness of their syntactic properties. In the idea that we have a "lexical storage" in our brain, thus, words are specified according to semantic, syntactic, orthographic, perceptual (pronunciation) and pragmatic attributes (Allport & Funnell, 1981; De Deyne et al., 2016). A prominent illustration of that is how phonological, orthographic, syntactic, and semantic information about words can be independently affected by damage to the brain (Emmorey & Fromkin, 1988).

Models of semantic representation

While there is no definitive proof of a concrete "lexical storage system" in the brain, scholars of different disciplines, cognitive scientists, neuroscientists, and linguists attempt to model how words are represented and retrieved. Several theoretical frameworks attempt to

explain how word meanings are represented and related to each other. These can be broadly categorized into attributional, distributional, and hybrid theories (Meteyard & Vigliocco, 2018). I will also refer to them as the theories or models of semantic representation.

Attributional Theories

According to these, a lexical representation of a concept is formed by bonding together different meaning constituents, features and attributes of a concept, e.g. chair – "wooden", "has legs." (Collins & Quillian, 1969; Meteyard & Vigliocco, 2018). Aitchison (2012a) states that there are perspectives that fall into the "atomic globule" viewpoint, whereupon "words are built up from a common pool of 'meaning atoms', with related words having atoms in common", which corresponds to the feature-based perspective. Purely attributional models face challenges in representing more abstract concepts adequately (Reggin et al., 2021).

Distributional Theories

In contrast, distributional theories propose that a word's meaning is directly derived from the immediate contexts in which it appears (Harris, 1954, as cited by De Deyne et al., 2016). The idea that distributional similarity and similarity in meaning highly correlate is called Distributional Hypothesis (Sahlgren, 2008).

These models, ranging from Latent Semantic Analysis (LSA) to modern word embeddings (like *word2vec*), have proven successful in capturing semantic similarity and relatedness (Baroni et al., 2014, cited by De Deyne & Perfors, 2016). However, they struggle to capture nuances not explicitly stated in text (for example, such models would fail at characterizing a banana as yellow as this information is not encoded in the text (Reggin et al.ha, 2021)) and therefore might be less effective than models based on human associations for certain tasks (Kumar et al., 2021;

De Deyne et al., 2016). Recent studies suggest that association-based models better predict diverse word properties (Vankrunkelsven et al., 2018) human similarity judgement (De Deyne et al., 2016b) or player performance in word guessing games (Kumar et al., 2021). The superiority of association-based models over text-based distributional models lends credence to the idea that mental representations integrate more than just statistical patterns.

Hybrid Theories

While many of the studies explore either side of the coin, some scholars argue (Andrews et al., 2005; Andrews et al., 2014) for the reconciliation and an inclusive view upon recognizing the limitations of purely attributional or distributional accounts. Hybrid theories propose that lexical knowledge arises from an integration of both feature-based and context-based information (Andrews, Frank, & Vigliocco, 2014). Aitchison's (2012a) "colweb" analogy captures this idea, suggesting words are linked through various relationships, including semantic similarity and co-occurrence.

Andrews et al. (2005) - both give complementary information about meaning

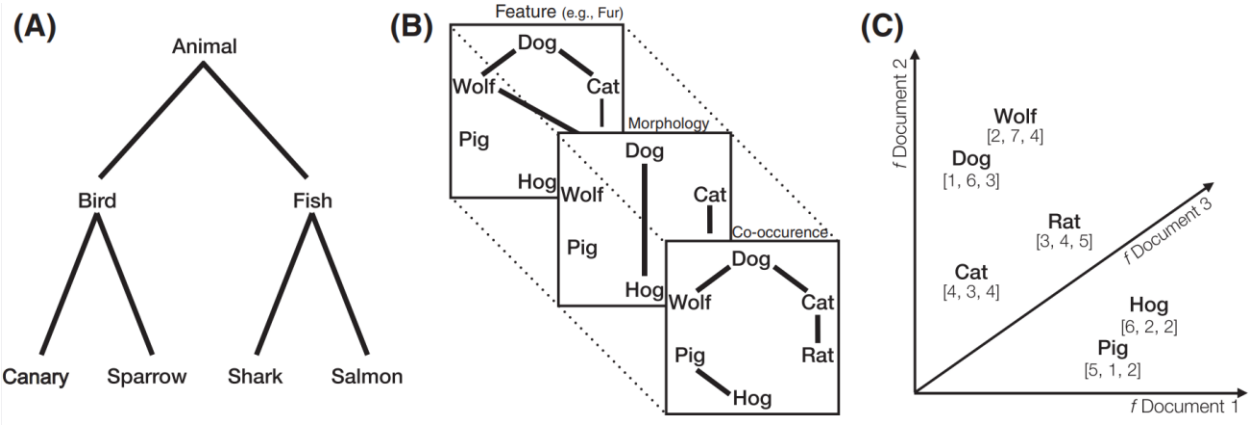
Hybrid perspective aligns well with network models of the mental lexicon, which can naturally represent multiple types of relationships between words (De Deyne et al., 2016a; Stella et al., 2018; Wulff et al., 2019). Such models are often considered more psychologically plausible as they can potentially capture the multifaceted nature of word meaning, including abstract concepts which may rely more heavily on linguistic context and affective factors (valency) (Reggin et al., 2021). The figure by Wulff et al. (2019) below illustrates computational models that align with the theories subject of this study:

- Semantic network theory (A) aligns with attributional theories.

- Vector-space models (C)such as WordNet correspond to distributional theories.
- Connectionist models (B), which utilize neural networks, reflect hybrid organization of mental lexicon.

Figure 1

Computational models of semantic representation



From Wulff et al. (2019) "New Perspectives on the Aging Lexicon"

Conceptualizing the mental lexicon as a network, where nodes represent words or concepts and edges represent the relationships between them as in B, has become increasingly influential (Stella et al., 2018; De Deyne et al., 2016a and 2016b; Beckage & Colunga, 2015; Tamariz-Martel, 2004). Recognizing that words are connected in multiple ways (semantically, phonologically, etc.), these models represent different relationship types in separate, interacting layers (Stella et al., 2018; Wulff et al., 2019). This allows for a more nuanced representation of the multifaceted nature of lexical knowledge.

Another important property is assortativity, the tendency for nodes to connect to other nodes

with similar properties. Finally, network analyses often reveal community structures – groups of words that are more densely connected to each other than to the rest of the network (De Deyne et al., 2016; Stella et al., 2018). These communities often correspond to thematic clusters, suggesting principles of organization beyond taxonomy (De Deyne et al., 2016). Stella et al. 's (2018) study is particularly notable in this regard: they found a "Largest Viable Cluster" (LVC) of words within the multiplex network. This cluster consists of words that are highly interconnected across all layers and removing the words within this cluster was found to negatively affect the imitation of processing for the model.

Studying Lexical Retrieval and Production

Various experimental paradigms are used to investigate how lexical knowledge is accessed and utilized. Methods include semantic priming, lexical decision tasks, word association tests and analysis of large-scale word association datasets and corpora (Pranoto & Afrilita, 2019; Beckage et al., 2015). While valuable for control, these tasks may lack ecological validity. Game-based methodologies are emerging as a way to study language use in more naturalistic, interactive, and goal-oriented settings (Allen et al., 2024; Kumar et al., 2021).

Gaps and Purpose

Primarily, what I found is that there is less understanding of how speakers *dynamically* choose retrieval strategies in real-time communication, especially under pressure. While factors like word frequency, concreteness, or part-of-speech were shown to influence the processing and retrieval (Brysbaert, 2018; Jessen et al., 2000; Sadoski et al., 1995), it is compelling to explore how these effects will manifest themselves in a naturalistic context.

Thus, this study aims to address these gaps by employing the Alias game as a naturalistic

paradigm and a mixed-method analysis of the game conversation data. I investigate the linguistic strategies players use to help partners guess target words, focusing specifically on how the choice and success of these strategies is impacted by the following factors: frequency, concreteness, and part-of-speech.

I aim at answering the following questions:

- What kind of linguistic strategies do players rely on, attempting to help their partners guess words?
- What are the factors that influence their choice of strategy, and factors that influence listeners' ability to guess the words quickly?
- How may the players' gameplay behavior reflect their mental lexicon organization?

The main hypothesis is as follows:

- Players will use different lexical strategies depending on word type (concrete or abstract), part of speech, frequency of the target word.
- Both attributional and distributional theories will be equally represented in these strategies, varying based on the factors listed above

This research offers methodological contributions to psycholinguistics by exploring the nature of real-time lexical access. For example, if some words are more frequently associated with a frequent collocation or an idiomatic expression and the corresponding association is produced during the game, it may be possible to observe some possible patterns of how lexical items are stored in the human brain.

Thus, the study hopes to contribute to the refinement of the current models of mental lexicon and language processing theories. Besides, insights into how humans rapidly retrieve and organize words can inform the design of more sophisticated language models and AI systems.

Methodology

Study Design and Rationale

While the encountered psycholinguistic studies on mental lexicon rely on controlled experiments such as Word association tests, this study examines natural language use in a social, collaborative setting with little interference from the researcher's side. I adopt a game-based research design, whereupon participants freely play an Alias game by themselves in groups of 4 or 6 using cards designed for this study. Allen et al. (2024) states that the game-based research design has a number of advantages – games are intuitive to play, are intrinsically motivating and rewarding. Since Alias is an interactional and communicative game, this approach allows for the observation of interactional dynamics with minimal artificial structuring.

Participants

The study targeted the general public aged 18-28 who are fluent speakers of Russian. An average participant is a Kazakh-Russian bilingual speaker between 18 and 28 years old interested in board games, as the participants were recruited in Astana, Kazakhstan. All participants were members of the Nazarbayev University community, as it is the closest place where board game playing sessions are organized regularly, but this population was not of a specific interest for this study.

I arranged 2 game sessions, each involving 4 participants (2 speaker-listener pairs) playing 6 rounds of the game. In each round, the speakers were given a card containing 8 words to be explained and guessed within a minute.

There were several participant recruitment channels. First, participants were recruited online by distributing recruitment messages in board game communities, such as NU Board Games. Second, a few participants were recruited on-site.

The participants of the two recordings are 8 Nazarbayev University students, with 3 pairs of strangers and 1 pair of friends participating.

Stimuli design

The stimuli set consists of 96 words distributed across 12 cards, with each card containing 8 words. These numbers were carefully chosen to ensure an even distribution of nouns, adjectives, and verbs (32 for each grammatical category), as well as account for concreteness/abstractness levels within the words, with 16 abstract and 16 concrete words in each grammatical category. Words that referred to tangible, perceivable objects, actions or qualities that can be directly experienced through the senses (such as *beaver*) were coded as concrete, while words that represented ideas, concepts, or emotions (such as *contagious*) rather than physical entities or actions.

Initially, I planned to use the cards from the original game pack, but since the distribution of parts of speech across the original pack was uneven, I had to randomly rearrange the words from the original game cards. Balancing for parts of speech, concreteness and word frequency facilitates a variety of communicative strategies during gameplay. Besides, it was necessary to prepare my own cards (Figure 1) to avoid bias in data and ensure consistency across categories.

Figure 2 The cards distributed to participants

Consistency was ensured in the following ways. First, each card was designed to contain a balanced mix of grammatical categories, with a minimum of 2 verbs, 2 nouns, and 2 adjectives, plus 2 additional words from any of these grammatical categories. Second, each card was balanced for concreteness, containing 4 abstract and 4 concrete words. Third, I also extracted the frequency of occurrence of each word from the National Corpus of Russian language to arrange words such that each card contains a mix of frequent, moderately frequent and infrequent words. The following table shows an example of one such card, illustrating the balance across grammatical categories, concreteness, and frequency:

Table 1

Sample stimuli card

#	Word	POS	Concreteness	Frequency
1	**бунтовать**	**V**	**A**	**2427**
2	охотиться	V	C	4270
3	коварный	A	A	28424
4	глубокий	A	C	58036
5	Пупок	N	A	870
6	жажда	N	C	10411

7	безумный	A	A	11094
8	тормозить	V	C	2815

Data collection procedures

The game sessions took place in public places, such as local coffee shops, open spaces at Nazarbayev University campus. Each session involved 4 participants completing six rounds of the game, during which their interactions are audio-recorded with minimal intervention. The pair members were seated next to each other against the opponent team, with the recorder located between them. Each team was given a predefined set of cards to use which they were not allowed to shuffle. The researcher was present, but did not interfere in participants' interaction except helping players with tracking time for each round.

Recordings captured players' spontaneous responses and word association strategies. For Nazarbayev University students, data collection sessions were conducted in public spaces on campus to maintain a familiar environment for participants and encourage natural language use.

The participants were not allowed to use the word of the same root or the translation of the target word in their explanation. In each round, each pair strived to explain and guess 8 words in any order within 1 minute duration. As the game progressed, teammates switched roles with each other to increase the ecological validity of the experiment. The participants were also free to shuffle the words within the card and skip some of them.

Data analysis procedures

Each game session was first transcribed by a Whisper AI model automatic transcription tool and proofread for mistakes. The transcripts were then reviewed manually to correct any inaccuracies. Then, each line of the game round was coded for the associative strategy type used

by the speaker. I also recorded the number of trials that took participants to guess (or not guess) the word. Below is a more detailed examination of each variable.

Variables

- Strategy. Categorical variable with 8 levels.
- Word Concreteness: Binary categorical variable coded as concrete (e.g., apple, run) or abstract (e.g., freedom, think) (coding procedure described earlier). It was relatively uniformly distributed.
- Target Word's Frequency: Discrete numerical variable extracted from the National Corpus of Russian Language (RusCorpora). The distribution followed Zipf's law. Frequency range is between 300 and ~80000 occurrences.
- Target Word's Part of Speech: Categorical variable with 3 levels (noun, verb, adjective). It was also relatively uniformly distributed, with some dispreference towards adjectives. Overall, players did not exhibit significant preference towards particular parts of speech.
- Trial & Guess**:** I considered an attempt to guess or a 0.5 sec pause as one trial. There were no instances where the player took long pauses causing the other one to change strategies; rather, all players tried to provide as many guesses as possible which resulted in multiple instances where one clue triggered a sequence of 2-4 guesses. As the number of times a speaker employs a strategy isn't always equal to the number of guesses provided by the listener, I preliminarily decided to only record those trials that corresponded to some strategy use.

Independent Variables under examination were target word's concreteness, part of speech and frequency of occurrence.

Due to use of multiple strategies for one word, several data points may correspond to only one target word. I decided to leave it as follows because otherwise I may have lost some important connections between strategy use.

While the coding of semantic fields should not present much question, the coding of collocations certainly does. Collocations were coded in the following way: once noticed, the target word was looked up in SKELL corpus that generates lists of strong collocates to the target word. If the target word and the speaker's explanation had a high collocation strength, the strategy used was coded as collocation.

Statistical test

Random Forest Model in R randomForest package was used to assess the impact of predictors on the likelihood of using specific strategies. This model was chosen because it can handle the interaction between a large number of possible predictors, and narrows down the focus to the most important ones. But the main reason is that the data is linearly dependent and there is a fair possibility of collinearity between the variables. Besides, the unbalancedness of data yields the use of the random forests model.

To account for possible interactions between the predictors, I also ran a random forest model using the randomForestSRC package which has a specific find.interaction function that calculates the strength of pairs of different predictors.

Further, I ran two multinomial regression models with the independent variables identified as most significant by the random forest model.

Results

Skipped words

It was natural to expect that not all words would be present in the game, because the design implied a one-minute time limit for the players to explain the words. However, what is interesting is players' *decisions* to skip some words.

At the beginning of each round, there was a brief time gap in which the speaker-player was silent. During this gap, each participant (all of them were in the role of a speaker 3 times) spent the time to examine the card. The players were not limited in their ability to skip or rearrange the order of the words, and in fact, in 5 rounds out of 6 conducted in session 1, the first described word was not a word that was first in the original order.

I presume that such patterns show deliberateness and purpose behind the players' word skipping pattern. While the lack of explanation for some words may also be attributed to the time constraints, it is also possible that some words in Table 2 were deprioritized by the speakers.

Table 2

Words skipped by players in BOTH game sessions

#	Word	Translation	POS	Frequency	Concreteness
1	удивительный	amazing	ADJ	27819	Abstract
2	управлять	to manage, rule	VERB	19705	Abstract
3	искренний	sincere	ADJ	14085	Abstract
4	наказать	to punish	VERB	9672	Abstract
5	типичный	typical	ADJ	9075	Abstract
6	преувеличивать	to exaggerate	VERB	2788	Abstract
7	искажение	distortion	NOUN	2547	Abstract
8	утонченный	refined	ADJ	2152	Abstract
9	самоуважение	self-esteem	NOUN	372	Abstract
10	прямой	straight	ADJ	42236	Concrete
11	жирный	fat	ADJ	8853	Concrete
12	обильный	abundant	ADJ	5859	Concrete

13	бородатый	bearded	ADJ	4282	Concrete
14	шершавый	rough	ADJ	2009	Concrete
15	шило	awl	NOUN	1137	Concrete

Out of 96 target words, 22 words were skipped by players of session 1, and 29 words were skipped by players of session 2. Table 2 presents only those words that were skipped by both groups.

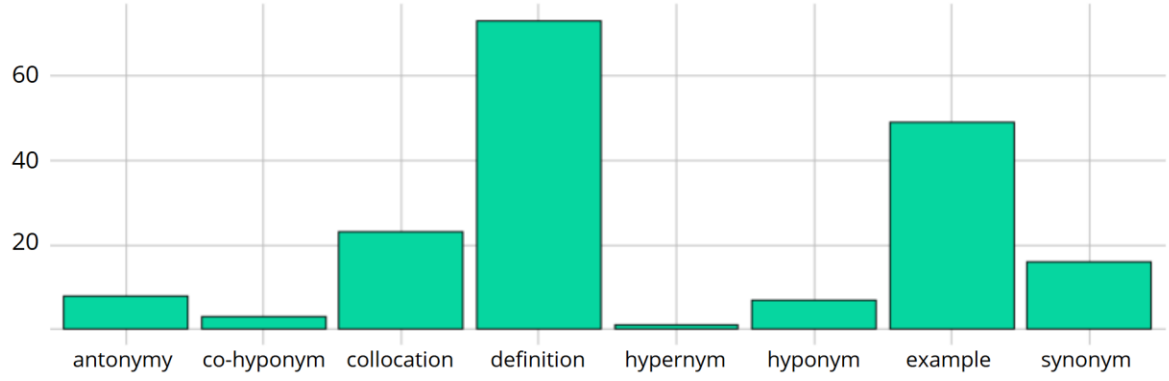
It is evident from the table that abstract words were skipped a bit more frequently (9 times over 6 times) than the concrete words, which in itself is an inconclusive result. However, what is common among these concrete words is that most of them (5 out of 6) were adjectives or words with lower frequency of occurrence. Most of the skipped words had a frequency of occurrence less than 10000. Low-frequency concrete adjectives and abstract words of all parts of speech were more likely to be skipped. It suggests that frequency influences both guessing success and players' willingness to attempt explanations. Adjectives were disproportionately represented among skipped words (9/15), possibly because they often require more nuanced explanations compared to nouns or verbs. For example, describing "шершавый" (rough) may require more creative or context-specific clues, which players may have avoided due to time pressure.

Patterns in Production

In lexical production, there was a slight preference for attributional strategies (definitions and semantic relations like antonyms, synonyms, hyponyms), though distributional strategies (collocations, real-life examples) were also prominent.

Figure 3

Frequency of use of different strategies

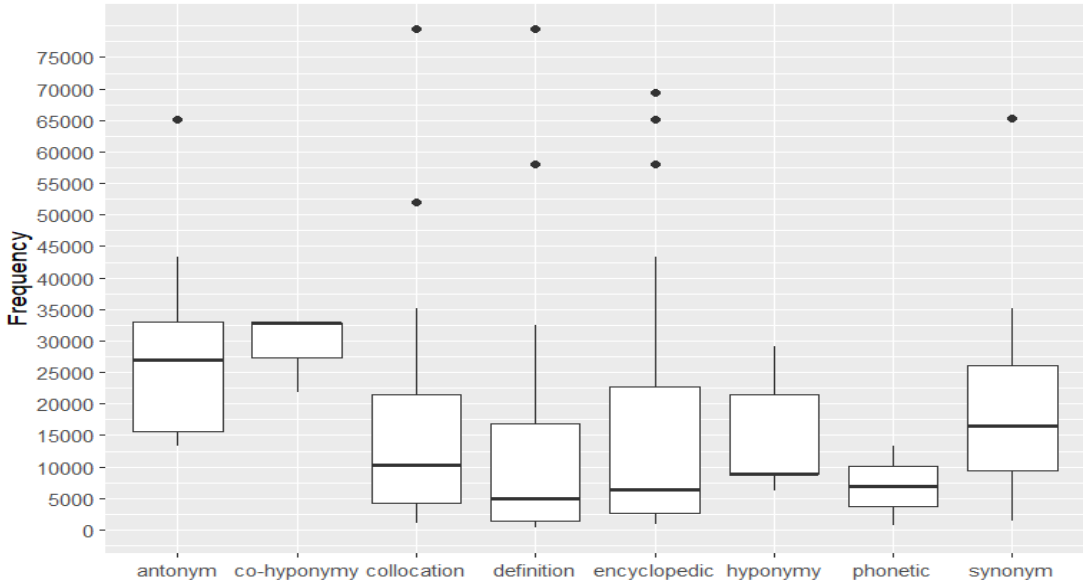


Further, I will talk about how each independent variable interacted with the strategy choice. There is ground to suggest that the word frequency and the combination of other variables may have some effect on the players' strategy choice.

First, let's discuss the possible effect of frequency. Figure 3 illustrates the box plot of each associative strategy and the corresponding frequency of the target words that were described using the said strategies:

Figure 4

Box plot of the strategy vs. the frequency of the target word



As displayed on the box plot, the words described by definitions and examples tended to be lower-frequency words: the median word frequency for both definitional and example-based strategies was around 5000. It means that 50% of the words that were described using these strategies had a word frequency lower than 5000. On the contrary, higher-frequency (>20000) words tended to be associated with relational strategies like antonymy, synonymy, and co-hyponymy.

With regard to the distinction of concreteness, however, no pattern was observed. Each of the 8 strategies was used in an equal proportion to describe concrete and abstract words.

Interaction between POS and strategy choice elicited only one distinct pattern: adjectives, unlike nouns or verbs, were more frequently explained using antonyms and synonyms.

Players' explanations

Now, let's draw the picture of what data was like: this section presents some player explanations from the games. Let's first consider the explanations built on semantic relations.

Synonymy (16 instances) was most frequent semantic relation present across 2 games, and was usually exemplified by a pattern "(it is) *like synonym"* such as below:

- *Типа натуральный ("It's like natural" as in "natural sciences")*

- *Органичный, органический? (естественный) ("Organic, organical?")*

Among 16 such instances, 8 attempted to explain *adjective* target words.

Antonymy-based (8) explanations also followed a consistent scheme of *A but the opposite:*

- *Поздний наоборот? (ранний)*

- *It is "late" but the opposite? (early)*

In 7 out of 8 cases, antonymy-based explanations were used for adjectives as well.

In explanations where **hyponymic relations (11)** were utilized, I observed 3 distinct ways it manifested:

1. Co-hyponymy (3 instances) - the player lists the members of a category to elicit the next member of the same category:
   - ➔ *Есть река, озеро, а большое,самое большое? (океан)*
   - ➔ *There is a river, a lake, and what is the biggest one? (ocean)*
2. Hyponymy (7) - listing members of the category to elicit the superordinate term:
   - ➔ *Ислам, христианство, буддизм. Что это все? (религия)*
   - ➔ *Islam, christianity, buddhism - what is that? (religion)*

   This strategy was used only for nouns.
3. Hypernym (1) - eliciting the subordinate term through providing the superordinate:
   - ➔ *Хобби, увлечение. Как бы один из… (творчество)*
   - ➔ *Hobby, interest, it's like one in this group… (art)*

Across all these instances, the player mentions the members of a particular semantic field and tries to give some direction to orient within the field (e.g. *made from stone but of the worse material*).

Definitional strategy involved the presentation of multiple attributes of a given word in a dictionary-definition format:

- ➔ *Такс, животное такое, у него зубы большие. Два зуба, он строит плотину* ***(бобёр)***
- ➔ *"It's an animal with big teeth. With two teeth, it builds a dam" (beaver)*

Collocation:

- ➔ *Ты можешь оказывать на друзей какое-то... (влияние)*
- ➔ *You can exert some … on your friends (influence)*

In **real-life example strategy**, the players tried to explain the target word by providing a detailed description of the linguistic context or a memorable referent:

- ➔ *Скажи, вот когда пацан не зовет девушку на свидание он... (тормоз, тормозить)*
- ➔ *"Tell me, when a guy doesn't dare ask a girl out on a date, then he is..." (slow thinker)*

Relationship between games strategies and attributional/distributional patterns

Definitions, semantic relations like antonyms, synonyms, hyponyms fell into the category exhibiting mainly attributional patterns because of their focus on associated features of the target word, while collocations and real-life examples aligned more with a distributional account due to their focus on linguistic contexts.

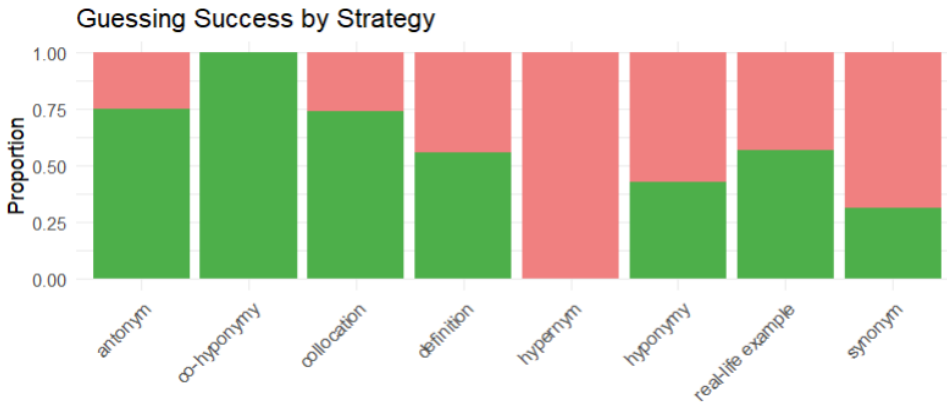
Patterns in Comprehension (success ratings)

This section will exemplify some of the important patterns related to the players' guessing success. Overall, the most effective strategies were co-hyponymy (100 % success rate),

antonymy and collocations (75%). Synonymy was the least effective (31%), while real-life examples and definitions showed moderate success (56-57%.)

Figure 5

Success rate by strategy



Now let's consider how the guessing success may differ across the factors such as frequency, part of speech and concreteness.

First, higher-frequency words were more likely to be guessed correctly (table below), with a mean frequency of 16,698 for guessed words compared to 10,670 for unguessed words. This is consistent with the well-documented frequency effect in psycholinguistics, where high-frequency words are processed more quickly and accurately due to their stronger representation in the mental lexicon.

Table 3

Word frequency across strategies and guessed/not guessed target words

Strategy	#	Success	Mean frequency			Median frequency		
			Overall	Not guessed	Guessed	Overall	Not guessed	Guessed

Antonymy	8	6 (75%)	29537	`20838	32437	26900	20838	29514
Synonymy	16	5 (31%)	19087	14971	34935	13251	12173	26172
Co-hyponymy	3	3 (100%)	29513	---	29513	32775	---	32775
Hyponymy	7	3 (43%)	15014	13892	21472	8931	8931	21472
Hypernym	1	0 (0%)	29138	29138	---	29138	29138	---
Collocation	23	17 (74%)	16867	14184	15730	10375	15135	6340
Definition	73	41 (56%)	10187	7663	9054	5658	6741	8568
Example	49	28 (57%)	13681	13036	11376	6428	6785	4914

Concreteness also highlighted certain patterns:

Table 4

Success rates by concreteness, trial order and corresponding word frequency

Category	**Guessed**	**1st Trial**	**2nd+ Trial**	**Mean Freq (Unguessed)**	**Median Freq (Unguessed)**	**Mean Freq (Guessed)**	**Median Freq (Guessed)**
Abstract	51%	66%	34%	12,495	11,094	19,449	13,165
Concrete	62%	65%	35%	8,003	3,747	14,612	6,073

First, it is noticeable that abstract words had a relatively low guessing success (51%).

Further, The mean frequency of guessed abstract words (19,449) was significantly higher than

that of unguessed abstract words (12,495). This suggests that even for abstract concepts, familiarity (as reflected by frequency) facilitated faster retrieval and guessing.

Similarly, the mean frequency of guessed concrete words (14,612) was higher than that of unguessed concrete words (8,003). However, the difference was less pronounced, indicating that concreteness may partially compensate for low frequency in facilitating word retrieval.

Overall, these results support distributional theories of semantic representation. However, what is also notable is that both the mean and median frequencies of words described with these strategies were significantly higher than those described using definitions or contextual and cultural knowledge. Low-frequency words may lack strong co-occurrence data (weak in distributional networks), but they still have attributes (visual, functional, encyclopedic). Even if a word is low-frequency, if the speaker can describe its distinctive features, and the listener shares that knowledge, it may still be guessed. Word frequency plays a weaker role in the success of attributional strategies (such as definitions and encyclopedic clues). Therefore, strong co-occurrence associations and presence of distinct notable attributes can override the effect of individual word frequency, suggesting the prevalence of a hybrid structure of the mental lexicon.

Table 5 presents the interaction of both concreteness and POS across strategies.

Table 5.

Success rates of strategies broken down by parts of speech and concreteness

Strategy	**Abstract**			**Concrete**		
	Noun	**Adjective**	**Verb**	**Noun**	**Adjective**	**Verb**
Antonymy	100%(1)	0%(2)	-	-	100%	-
Synonymy	100%(1)	0%	-	50%	0%	67% (3)
Co-hypon	-	-	-	100%	100%	-

Hyponymy	75%	-	-	0%	-	
Collocation	57%	-	100%	80%	100% (2)	56%
Definition	64%	60%	33%	69%	60%	70%
Encyc	60%	56%	56%	100%	29%	62%

Note: ignore the grey cells

Antonymy was highly effective for concrete adjectives (100%), but completely failed for abstract adjectives (0%). Synonymy was completely ineffective for adjectives, failing to yield a single correct guess regardless of concreteness. However, it was moderately successful for concrete verbs (67%). These results suggest that synonymy overall may be an unstable retrieval strategy for comprehension, as synonyms often have subtle differences in meaning that can lead to confusion rather than direct word activation. The complete fail with regard to adjectives is notable: while more than half of the words described by synonymy were adjectives, it shows that there is an asymmetry between language production and language comprehension: while retrieval through direct substitution is easy, it is unreliable, likely because of overlapping but non-identical meanings which confuses the interlocutor who interprets the guess.

Collocation was among the most effective strategies overall, performing particularly well for concrete adjectives (100%), concrete nouns (80%), and abstract verbs (100%). These results suggest that frequently occurring multi-word expressions aid lexical retrieval significantly, especially when a target word is embedded within a common phrase. However, collocations were slightly less effective for concrete verbs (56%), indicating that not all verbs are equally represented in fixed expressions.

Definition-based strategies showed moderate to high effectiveness across most categories, particularly for concrete verbs (70%) and concrete nouns (69%). However, their

success rate dropped for abstract verbs (33%), suggesting that defining abstract actions is more difficult than defining tangible actions. The moderate success for abstract adjectives (60%) and concrete adjectives (60%) indicates that definitional strategies can be useful but are not always the most efficient retrieval method.

Encyclopedic knowledge was highly effective for concrete nouns (100%), reinforcing the idea that world knowledge and shared cultural references play a significant role in retrieving tangible entities. However, it was far less effective for concrete adjectives (29%), suggesting that adjectives may rely more on relational or comparative structures than on encyclopedic associations. For abstract nouns (60%), abstract adjectives (56%), and abstract verbs (56%), encyclopedic knowledge remained moderately useful, but its success rate was not as high as for concrete nouns.

Overall, this data provides strong evidence that lexical retrieval is influenced by both attributional (feature-based) and distributional (co-occurrence-based) factors, with different strategies proving effective depending on the type of word being retrieved.

Results of the statistical tests

Identifying variable importance and most important pattern through Random Forest

Table 6

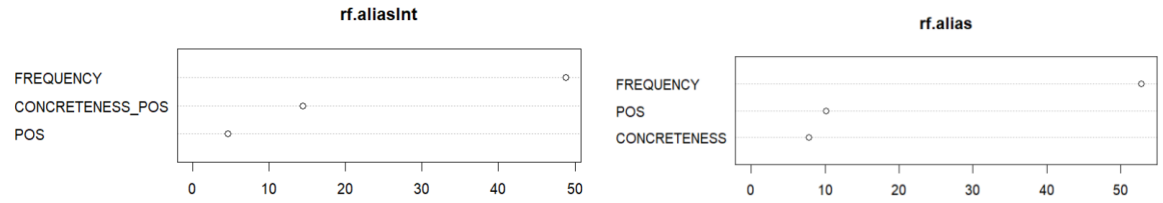
Interaction strength between predictors

	Frequency	**POS**	**Concreteness**
Frequency	*0.07*	*0.33*	*0.32*
POS	*0.15*	*0.09*	*0.24*
Concreteness	*0.22*	*0.53*	*0.26*

Note: *0 means no interaction, 1 means very strong interaction.*

Figure 8

Variable importance plots for models with (1) and without (2) variable interaction

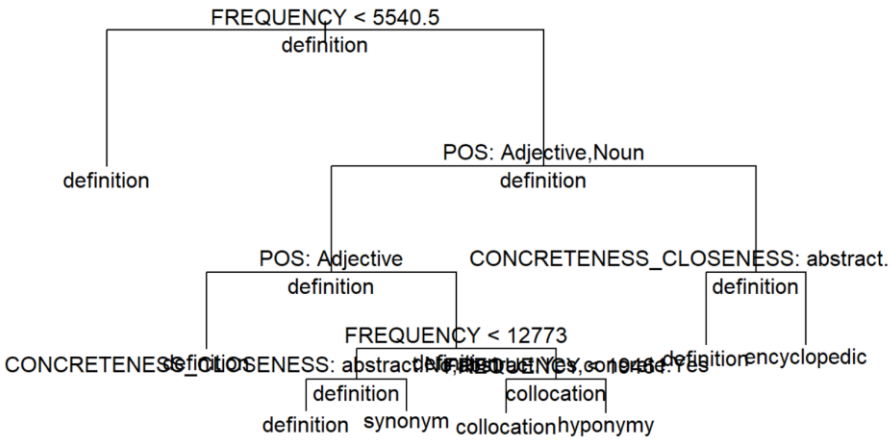


While concreteness and POS did not exemplify any patterns by themselves, based on their interaction was identified as having stronger effect together, rather than separately: Therefore, it was decided to validate the model that considered the interaction of these predictors using the decision tree model. The error rate for the latter was ~52% compared to the ~49% in the former model. To validate the model, I created a decision tree training it on 75% of my data and tested the tree on the other 25%. The train and test split was done randomly. The accuracy of the decision tree model on the test data was 50%.

Because the tree looked very complex and difficult to interpret due to too many nodes, I pruned the model (reduced the complexity of the tree by keeping only the most important nodes). The pruned model exhibited the accuracy of 60%.

Figure 9

Pruned decision tree



Thus the model best predicts the use of a definitional strategy, and the decision trees also show that the model predicts frequency as the variable with most significant effect: the model predicts that the words with frequency less than 5540 are most likely to be explained using a definitional strategy. This finding supports the previous idea that lower-frequency words tend to be characterized more by their attributes.

Discussion

Recap of key findings

Our analysis of the Alias gameplay data revealed several key findings. First, players employed a diverse range of strategies, with definitional and example-based clues being the most frequent overall, but with significant use of strategies based on semantic relations (synonymy, antonymy, hyponymy) and collocations. Most crucially, the choice of strategy was to a moderate extent modulated by the properties of the target word. Word frequency emerged as a particularly strong predictor: low-frequency words predominantly elicited definitions or example-based knowledge, whereas high-frequency words were more likely to be described using semantic relations or collocations. Statistical analysis confirmed a moderate interaction between concreteness and part-of-speech influencing strategy choice.

Strategy success rates varied; collocations and antonymy (especially for concrete adjectives) proved highly effective, while synonymy was notably less successful, particularly for adjectives. Definitional strategies showed moderate success rates and appeared less dependent on word frequency (Table). Finally, players tended to skip words that were abstract (especially adjectives and low-frequency verbs) and guessed them with a lower (50%) success rate. This points us towards a possible inherent difficulty in concisely describing concepts that lack easily accessible, stable features or strong, unique associations, especially under time pressure (Reggin et al., 2021).

Drawing a relation to the hypotheses

These findings provide substantial support for my central hypotheses. Moderately, the strategy choice seems to be modulated by word frequency, concreteness, and POS, which aligns with Kumar et al.'s (2021) game based study findings. It confirms that lexical retrieval is not a uniform process but is adapted to the characteristics of the target word. The co-existence and differential use of both feature-based and co-occurrence-based strategies supports the hypothesis that the mental lexicon integrates both attributional and distributional information.

Particularly, the frequency effect requires closer examination. The reliance on definitions for low-frequency words suggests that when contextual exposure is limited, resulting in potentially weaker or less accessible storage patterns, speakers default to describing core semantic features. Conversely, the preference for relational and collocation strategies for high-frequency words indicates that repeated exposure strengthens associative pathways, making the plays choose more efficient and riskier strategies (De Deyne et al., 2016). This aligns with the concept of high-frequency words acting as central hubs in semantic networks (Stella et al., 2018), possessing numerous, easily activated connections.

The differential success rates offer insights into retrieval efficiency and cue diagnosticity. The high success of collocations and antonyms suggests they provide constrained, highly predictive cues for the listener. Collocations employ highly probable linguistic sequences, while antonyms offer clear semantic contrasts. The striking failure of synonymy for adjectives might stem from the nuanced, context-dependent nature of adjectival meaning; providing a near-synonym might introduce more ambiguity than clarity for the listener, highlighting a potential asymmetry between ease of production for the speaker and ease of comprehension for the listener. The moderate success of definitions, relatively independent of frequency, suggests that accessing core features can be a robust, but potentially slower, strategy available even for less common words.

Finally, the pattern of skipped words – primarily abstract items and low-frequency concrete adjectives – points towards the inherent difficulty in concisely describing concepts that lack easily accessible, stable features or strong, unique associations, especially under time pressure (Reggin et al., 2021).

These results offer empirical support for hybrid theories of the mental lexicon (Aitchison, 2012; Andrews et al., 2014). The fluid switching between strategies that access semantic features (definitions) and those that access usage patterns (collocations, relations), contingent on word frequency, demonstrates an integrated system rather than one reliant solely on attributes or distribution. This aligns with multi-layer network proposals (Stella et al., 2018; De Deyne et al., 2016), where different strategies might tap into distinct but interconnected layers of representation: a feature-based layer and a layer that stores information about linguistic contexts. Furthermore, the observation that strategies like definitions often implicitly combine feature knowledge (e.g., 'beaver' *is an animal*) with usage/contextual knowledge (*builds dams*), blurs the

strict boundary between attributional and distributional strategies.

Comparing our findings with network model concepts, the effectiveness of collocations for frequent words strongly supports the idea of robust associative links (high edge weights) between frequently co-occurring words in semantic networks (De Deyne et al., 2016; Beckage & Colunga, 2015). The frequency effect mirrors the role of hubs (high-frequency words) in network topology, which are thought to facilitate faster spreading activation or easier access to diverse neighbours (Steyvers & Tenenbaum, 2005, cited by Beckage & Colunga, 2015; Stella et al., 2018).

Contrasting our results with word association data, the Alias task elicits more complex, goal-directed strategies than simple free association as in Pranoto & Afrilita's (2019) study. Players select clues strategically, not just based on raw associative strength, but on perceived utility for the listener; thus the associations are communication-oriented. This highlights the difference between probing the structure of I-language (associative patterns over distributional) (as association tasks do, De Deyne & Perfors, 2016; Vankrunkelsven et al., 2018) and observing how I-language is used strategically in a communicative, time-pressured context. Compared to other game-based studies like Kumar et al. (2021), these results expand our understanding in the following direction: while associative knowledge might be a strong predictor overall, speakers flexibly draw on both associative/distributional and feature-based knowledge depending on the specific word and context.

Nuances behind the classification of strategies

It is important to note, the distinction between attributional and distributional strategies is not always clear-cut, particularly with regard to these two strategies: while definitions were the most common strategy used by players, they cannot be classified as purely distributional or

purely attributional. Definitions often include both semantic features (attributional) and contextual usage patterns (distributional). For example, defining "бобёр" (beaver) as "an animal that builds dams" involves both the intrinsic semantic properties of the word (e.g., "animal," "builds dams") and its typical usage in natural language (e.g., the co-occurrence of "beaver" and "dam"). Similarly, some definitions contain synonymic relationships (e.g., "ударить" [to hit] defined as "бить" [to beat]). Such variability suggests that lexical organization is more nuanced than a strict binary between distributional vs. attributional strategies. Thus, the naturalistic game-based study design suggests a complex picture, where different storage patterns and pathways might overlap, rather than rigid categorical distinctions such as purely attributional or purely distributional.

Conclusion, implications and future research

This study provides evidence for the hybrid nature of the mental lexicon, demonstrating that word meanings are processed through a combination of feature-based and context-based mechanisms. Players dynamically adapt their descriptive strategies, employing a mix that reflects both the features (attributional information) and the usage patterns (distributional information) of target words. It carries several implications. Methodologically, it underscores the value of using naturalistic, interactive tasks like Alias to capture language processing and retrieval that might be obscured in more constrained experimental settings (Allen et al., 2024). Theoretically, it provides empirical support for hybrid models of the mental lexicon and demonstrates the adaptive nature of lexical retrieval (Tamariz-Martel, 2004).

The dataset was relatively small, and the unbalanced nature of the data impacted the reliability of the statistical testing and interpretations. Besides, the participant sample was homogeneous (Nazarbayev University students, Russian speakers), and thus lacks

generalizability to other populations and languages. Future research should aim to address these limitations by using larger, diverse datasets and participant samples across multiple languages. Besides, collecting listener-side data (guessing times, errors) could provide a richer picture about the comprehension of the explanations. Finally, a specific qualitative exploration of clause participants give about abstract words, informed by recent work (Reggin et al., 2021), would also be valuable.

References

Aitchison, J. (2012a). Word-webs: Semantic networks. In *Words in the mind: An introduction to the mental lexicon*. John Wiley & Sons.

Allen, K., Brändle, F., Botvinick, M., Fan, J. E., Gershman, S. J., Gopnik, A., ... & Schulz, E. (2024). Using games to understand the mind. *Nature Human Behaviour*, 1-9.

Allport, D. A., & Funnell, E. (1981). Components of the mental lexicon. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *295*(1077), 397-410.

Andrews, M., Vigilocco, G., & Vinson, D. B. (2005). The role of attributional and distributional information in semantic representation. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 27, No. 27).

Andrews, M., Frank, S., & Vigliocco, G. (2014). Reconciling embodied and distributional accounts of meaning in language. Topics in cognitive science, 6(3), 359-370.

Beckage, N., Aguilar, A., & Colunga, E. (2015). Modeling lexical acquisition through networks. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 37).

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. Current directions in psychological science, 27(1), 45-50.

Clark, H. H. (1970). Word associations and linguistic theory. *New horizons in linguistics*, *1*, 271-286.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior*, *8*(2), 240-247.

De Deyne, S., Kenett, Y. N., Anaki, D., & Faust, M. (2016a). Large-scale network representations of semantics in the mental lexicon. In *Big data in cognitive science* (pp. 183-189). Psychology Press.

De Deyne, S., Perfors, A., & Navarro, D. J. (2016b). Predicting human similarity judgments with distributional models: The value of word associations. In *Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers* (pp. 1861-1870).

Emmorey, K. D., & Fromkin, V. A. (1988). The mental lexicon. *Linguistics: the Cambridge survey*, *3*, 124-149.

Harris, Z. S. (1954). Distributional structure. *Word*, *10*(2-3), 146-162.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: the case of frequency of occurrence. American psychologist, 39(12), 1372.

Jessen, F., Heun, R., Erb, M., Granath, D. O., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The concreteness effect: Evidence for dual coding and context availability. Brain and language, 74(1), 103-112.

Kumar, A. A., Steyvers, M., & Balota, D. A. (2021). Semantic memory search and retrieval in a novel cooperative word game: A comparison of associative and distributional semantic models. *Cognitive Science*, *45*(10), e13053.

Meteyard, L., & Vigliocco, G. (2018). Lexico-semantics. *The Oxford handbook of psycholinguistics*, 71-90.

Pranoto, B. E., & Afrilita, L. K. (2019). The organization of words in mental lexicon: evidence from word association test. *Teknosastik*, *16*(1), 26-33.

Reggin, L. D., Muraki, E. J., & Pexman, P. M. (2021). Development of abstract word knowledge. *Frontiers in psychology*, *12*, 686478.

Riordan, B., & Jones, M. N. (2011). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science*, 3(2), 303-345.

Sadoski, M., Goetz, E. T., & Avila, E. (1995). Concreteness effects in text recall: Dual coding or context availability?. Reading Research Quarterly, 278-288.

Sahlgren, M. (2008). The distributional hypothesis. *Italian Journal of linguistics*, *20*, 33-53.

Stella, M., Beckage, N. M., Brede, M., & De Domenico, M. (2018). Multiplex model of mental lexicon reveals explosive learning in humans. *Scientific reports*, *8*(1), 2259.

Tamariz-Martel Mirêlis, M. (2004). Exploring the adaptive structure of the mental lexicon.

Vankrunkelsven, H., Verheyen, S., Storms, G., & De Deyne, S. (2018). Predicting lexical norms: A comparison between a word association model and text-based word co-occurrence models. *Journal of cognition*, *1*(1), 45.

Wulff, D. U., De Deyne, S., Jones, M. N., & Mata, R. (2019). New perspectives on the aging lexicon. *Trends in cognitive sciences*, *23*(8), 686-698.

Appendix

Link to the dataset:

https://github.com/anonforconf/patterns_lex_processing/tree/main