*Bilingualism: Language and Cognition*

cambridge.org/bil

# Contextual diversity and picture naming: The role of aging and bilingualism

Mikayla Trudeau-Meisner[1], Brendan T. Johns[2] and Vanessa Taler[1,3]

[1]School of Psychology, University of Ottawa, Ottawa, ON, Canada; [2]Department of Psychology, McGill University, Montreal, QC, Canada and [3]Bruyère Health Research Institute, Ottawa, ON, Canada

**Research Article**

**Cite this article:** Trudeau-Meisner, M., Johns, B.T. and Taler, V. (2025). Contextual diversity and picture naming: The role of aging and bilingualism. *Bilingualism: Language and Cognition* 1–17. https://doi.org/10.1017/S1366728925100229

Received: 9 June 2023
Revised: 12 May 2025
Accepted: 22 May 2025

**Keywords:**
lexical access; picture naming; contextual diversity; corpus-based models; aging; bilingualism

**Corresponding author:**
Vanessa Taler;
Email: vtaler@uottawa.ca

## Abstract

Word frequency has long been considered an essential aspect of psycholinguistic theory. However, research has shown that measures of contextual and semantic diversity provide a better fit to lexical decision and naming data than word frequency. The current study examines the role of contextual and semantic diversity in picture naming ability across aging and bilingualism. A picture naming experiment was conducted with six groups of participants: younger monolinguals, older monolinguals, younger L1 English bilinguals, older L1 English bilinguals, younger L2 English bilinguals and older L2 English bilinguals. Consistent with previous findings, the contextual diversity measure accounted for more variance in the picture naming data than word frequency. Furthermore, older adults and L1 English bilinguals were more sensitive to semantic diversity information, while younger adults and L2 English bilinguals relied more on age of acquisition in their lexical organization.

## 1. Introduction

Picture naming tasks are commonly used in psycholinguistics to examine language function because they tap into the structure and accessibility of the mental lexicon (Cuitiño et al., 2019). In these tasks, participants must say the word that corresponds to the image shown. Outcome measures include accuracy (i.e., the number of correct responses) and latency (i.e., the amount of time that elapses between stimulus onset and naming). Traditionally, naming latencies have been predicted by variables such as word frequency (WF) and age of acquisition (AoA). More recently, contextual diversity (CD), which is defined as the number of distinct linguistic contexts in which a word occurs, has emerged as an important factor in lexical processing (Adelman et al., 2006). This measure is grounded in the *principle of likely need* from the rational analysis of memory (Anderson, 1974; Anderson & Milson, 1989; Anderson & Schooler, 1991), according to which a word that has been encountered in many contexts is more likely to be needed in future contexts, and thus should be more readily accessible in the lexicon (see Jones et al., 2017 for a more thorough discussion of these issues). CD is measured by a document or context count, which is operationalized as the number of documents or contexts (with a context being defined at different lexical units; e.g., a sentence, paragraph or chapter in a book) in which a word occurs across a corpus (Johns et al., 2016b; Jones et al., 2012). Among others, Adelman et al. (2006) compared WF and CD and found that CD was more predictive of word naming and lexical decision reaction times than WF (Adelman & Brown, 2008; Brysbaert & New, 2009; Johns et al., 2016b, 2022).

### 1.1. *Effects of aging and bilingualism on lexical access*

The present study examines the role of CD in picture naming across aging and bilingualism. First, studies have shown that bilingualism affects picture naming ability. When compared to monolinguals, bilinguals are less accurate (Bialystok et al., 2008; Kohnert et al., 1998; Sheppard et al., 2016) and slower to respond (Gollan et al., 2005a, 2005b, 2007; Ivanova & Costa, 2008; Roberts et al., 2002). Bilingual adults also have more tip-of-the-tongue retrieval failures for object names than monolinguals (Gollan et al., 2005a). Similarly, findings regarding the impact of normal aging on picture naming ability point toward an age-related decline. Older adults are significantly less accurate (Ardila & Rosselli, 1989; Burke & Mackay, 1997; Feyereisen, 1997; Ivnik et al., 1995; Mackay et al., 2002; Mitrushina & Satz, 1995; Nicholas et al., 1989; Rosselli et al., 1990; Zec et al., 2005, 2007) and are slower to respond on picture naming tasks than younger adults (Paesen & Leijten, 2019; Shafto & Tyler, 2014). In addition, older adults also experience more tip-of-the-tongue retrieval failures than younger adults (Burke & Mackay, 1997; Shafto et al., 2007; Shafto & Tyler, 2014; Silagi et al., 2015).

Mägiste's (1979) interdependence hypothesis of bilingual storage suggests that multilinguals experience slower lexical retrieval due to less frequent use of their languages and interference from competing language systems. The *frequency lag hypothesis* (Gollan et al., 2005b, 2008, 2011) holds that the regular use of two languages creates a disadvantage because bilinguals have less



CAMBRIDGE UNIVERSITY PRESS

experience with words in each of their languages compared to monolingual speakers of that language. However, because bilinguals routinely navigate between two linguistic systems and must attend to language-specific cues, they may develop enhanced sensitivity to contextual information (Gollan & Ferreira, 2009; Johns et al., 2016b). This ability may enable them to better distinguish among competing lexical items by exploiting CD, or the variety of linguistic or situational contexts in which a word appears.

Similarly, poorer picture naming performance in older adults has been attributed to an age-related decline in the efficiency of lexical access (Burke & Shafto, 2004; Kavé et al., 2010). Ramscar et al. (2014) proposed the *information accumulation perspective on aging*, which suggests that the age-related decline in performance on cognitive tests reflects accumulated linguistic knowledge over the lifespan. In other words, older adults perform worse on cognitive tasks not because of cognitive decline, but because of the higher information-processing costs of navigating the cumulative knowledge in their cognitive systems (Ramscar et al., 2014). CD may mitigate this challenge by providing more retrieval cues and reducing reliance on frequency-based access. Words that have been experienced in more diverse contexts are more richly embedded in semantic memory, which could make them more accessible despite a general age-related decline in processing speed or control.

Both bilingualism and aging can impact language production. More specifically, they may affect lexical access, which involves a speaker activating and choosing the right word from their mental lexicon (Levelt et al., 1999). It has been hypothesized that both languages are activated when bilinguals produce language, and cognitive processes inhibit the nontarget language in language-specific situations (Costa & Caramazza, 1999; Green, 1998; Hermans et al., 1998). Initially, all activated options compete for selection, but a late-acting process reduces the activation of the nontarget language to enable the selection of an appropriate response (Misra et al., 2012). Therefore, the stronger first language (L1) must be inhibited to enable production in the weaker second language (L2), which in turn impacts performance in the L1 (Kroll et al., 2008; Levy et al., 2007; Linck et al., 2009; Philipp et al., 2007). Aging, in turn, is associated with declines in processing speed and inhibitory control (Christ et al., 2001; West & Alain, 2000), which can impair the suppression of competing words and increase the likelihood of retrieval difficulties, such as tip-of-the-tongue experiences (Burke & Mackay, 1997).

### *1.2. The role of contextual diversity and semantic distinctiveness*

CD may help in each of these cases by strengthening word representations through repeated exposure in different situations (Adelman et al., 2006; Johns et al., 2016b). Words that are encountered across diverse contexts may form stronger or more varied retrieval pathways, in turn making them easier to access in populations where retrieval requires more effort. As such, CD may serve as a compensatory mechanism that supports more efficient lemma selection, which could help bilinguals and older adults during lexical access.

While CD has emerged as an important factor in lexical processing, it cannot operate in isolation. Picture naming is also shaped by more established psycholinguistic variables, most notably age of acquisition and word frequency, both of which have long been recognized as influences on word retrieval. AoA refers to the age at which a given lexical item is learned. The consensus is that the earlier a word is learned, the faster and more accurately it can be accessed (Alario et al., 2004; Barry et al., 1997; Bonin et al., 2002, 2003; Cuetos et al., 1999; Dell'acqua et al., 2000; Ellis & Morrison, 1998; Khwaileh et al., 2018; Perret & Bonin, 2019; Snodgrass & Yuditsky, 1996; Valente et al., 2014). WF refers to the number of times that a word appears in a particular corpus (Perret & Bonin, 2019). High-frequency words will be accessed more quickly and accurately than lower-frequency words (Alario et al., 2004; Barry et al., 1997; Bonin et al., 2003; Cuetos et al., 1999; Cuitiño et al., 2019; Ellis & Morrison, 1998; Snodgrass & Yuditsky, 1996), although it should be noted that some studies have failed to replicate this effect (Bonin et al., 2002; Dell'acqua et al., 2000; Valente et al., 2014). WF, in particular, has been a central aspect of psycholinguistic theory for decades (Brysbaert et al., 2018) and is a key component in most models of lexical organization and word integration (Coltheart et al., 2001; Goldinger, 1998; Morton, 1969; Murray & Forster, 2004; Norris, 2006).

Research on the level of representation at which WF, AoA, and CD effects occur within the language production system is inconclusive. In a seminal paper, Jescheniak and Levelt (1994) proposed that the WF effect arises during the retrieval of word forms (also known as lexeme retrieval). However, a growing body of evidence suggests that WF affects lemma selection (Corps & Meyer, 2023; Navarrete et al., 2006; Wheeldon & Monsell, 1992), which is the process by which one selects the semantically appropriate item (Jescheniak & Levelt, 1994; Roelofs, 1992). In fact, Corps and Meyer (2023) determined that in picture naming tasks, WF does not exclusively affect word form retrieval, but also lemma access.

There is debate over whether the locus of AoA effects reflects internal properties of the lexicon or external properties of the learning process (Hernandez & Li, 2007). WF and AoA are correlated (Wang & Chen, 2020), and Brysbaert and Ghyselinck (2006) propose that AoA effects are partly frequency-related and partly frequency-independent. This frequency-independent component of the AoA effect has been thought to have a semantic locus (Brysbaert & Ellis, 2016; Wang et al., 2023). Wang and colleagues used event-related potential (ERP) techniques to examine the underlying mechanisms of the L2 AoA effect on three distinct levels: sub-lexical, lexical and semantic. First, they found that L2 AoA effects specifically were not influenced by WF (Wang et al., 2023). However, the authors also determined that the L2 AoA effect had both lexical and semantic routes and that, like WF, AoA effects arise from both semantic representation and spelling–sound connections (Wang et al., 2023).

CD and WF are correlated, which suggests that they may reflect the same underlying processes (Adelman et al., 2006; Vergara-Martínez et al., 2017). If this is the case, CD could substitute for WF in models of lexical organization and word integration with few theoretical implications (Plummer et al., 2014). However, ERP research shows that CD and WF originate from different sources during the access of lexical-semantic representations (Vergara-Martínez et al., 2017). Vergara-Martínez et al. (2017) determined that higher CD words elicit larger negativities than lower CD words, which is the opposite of WF, where higher frequency words elicit smaller negativities. Furthermore, they determined that the scalp distributions differ: the anterior distribution of the CD effect is consistent with previous effects related to semantically richer words (Vergara-Martínez et al., 2017). Overall, these findings suggest that the CD effect resembles other factors related to "semantic richness" (Rabovsky et al., 2012), but this effect cannot be explained in terms of these variables because the stimuli were matched for these factors (Vergara-Martínez et al., 2017).

Understanding the locus of effects for WF, AoA and CD helps clarify how these variables impact picture naming in bilinguals and older adults. Recent research suggests that WF and CD influence lexical access at the lemma level (e.g., Corps & Meyer, 2023; Vergara-Martínez et al., 2017). Factors that may impact lemma selection, including increased lexical competition in bilinguals (Kroll et al., 2008; Levy et al., 2007; Linck et al., 2009; Philipp et al., 2007) or decreased inhibitory control in older adults (Christ et al., 2001; West & Alain, 2000), may influence naming performance. Furthermore, AoA effects, which involve both lexical and semantic processes (Wang et al., 2023), may be amplified in bilinguals, who have less exposure to high-frequency words (Gollan et al., 2005b, 2008, 2011), and in older adults, who may rely more on semantic access as phonological retrieval declines with age (Burke & Shafto, 2004). CD's association with semantic richness (Rabovsky et al., 2012) aligns with evidence that bilinguals are attuned to contextual cues, suggesting that CD may support lexical access when retrieval is more effortful. Together, these findings highlight that lexical access is shaped not only by how often or how early a word is learned but also by the diversity of its conceptual and semantic representations.

The above sections focusing on the mechanisms of AoA, CD and WF highlight the importance of semantic information in models of lexical organization. These lexical characteristics not only capture how often a word is encountered or when it is learned but also reflect the diversity of the contexts in which it is used. Building on this view, Johns (2021; see also Chang et al., 2023 and Johns & Jones, 2022, for additional analyses) offers a refinement of traditional CD measures, using these more socially-based theoretical constructs. This proposal offers two theoretical notions of the types of contexts in which words can occur: 1) *discourse contextual diversity* (DCD) or 2) *user contextual diversity* (UCD). These measures were derived from analyzing communication patterns of over 300,000 users across more than 30,000 discourse topics (subreddits) on the internet forum Reddit, with a total word count exceeding 55 billion for each metric. The UCD measure quantified the number of users who used a particular word, while the DCD measure tracked the number of discourses in which the word appeared. These count measures significantly improved upon the traditional WF and CD metrics for both response time and accuracy data in both lexical decision and naming tasks (Chang et al., 2023; Johns, 2021; Johns & Jones, 2022), as well as related tasks such as item-level effects in recognition memory (Johns et al., 2022).

DCD and UCD measures offer substantial improvements over WF and CD measures, especially concerning accuracy (Johns, 2021). This finding suggests that measuring contextual word usage at the discourse and user level is more advantageous than using smaller, non-socially based, count measures.

### 1.3. Semantic distinctiveness model

Although CD measures consistently outperform WF measures (Adelman et al., 2006; Adelman & Brown, 2008; Brysbaert & New, 2009), there may still be an important information source missing: the semantic diversity (SD) of the contexts in which a word occurs (Johns et al., 2016b). To examine the role of SD in lexical organization, Jones et al. (2012) proposed the *semantic distinctiveness model.* The semantic distinctiveness model belongs to a class of models entitled distributional models of language, which use the statistical structure of the language environment to learn the meaning of words (see Kumar, 2021 for a recent review). However, instead of constructing the meaning of a word (which the model can do; Johns & Jones, 2008), the goal of the model is to generate more refined measures of a word's lexical strength. The strength measures generated by the semantic distinctiveness model are based on the semantic diversity of the contexts (with context being defined differently across the development of the model) in which a word occurs across a corpus, where words that occur in more unique semantic contexts have higher memory strength than those in more redundant contexts (Jones et al., 2012). The semantic distinctiveness model has been repeatedly demonstrated to provide a more accurate measure of a word's strength in memory, evaluated upon datasets using varied behavioral data from across language and episodic memory (Chang et al., 2023; Johns, 2021; Johns et al., 2016a, 2016b, 2020, 2022; Johns & Jones, 2022).

Transformations based on the semantic distinctiveness model have been used to modify the DCD and UCD measures to better explain the importance of the semantic content of linguistic contexts (Johns, 2021). Previous research indicates that SD-transformed models fit the relevant data better than count models (Chang et al., 2023; Johns, 2021; Johns et al., 2016a; 2020; Jones et al., 2012). Coherent with CD measures, the semantic distinctiveness model counts the number of contexts a word appears in. However, each context is given a graded measure between 0 and 1 depending on the semantic similarity between a word's representation (stored in memory) and the context representation. More surprising, or distinct, contextual usages of a word are given greater weight in a word's strength in memory. This is accomplished by taking the vector cosine between a word's representation stored in memory and the current context representation and modifying this with an exponential transformation where high similarity values are transformed to low SD values and low similarity values are transformed to high SD values (Jones et al., 2012).

The main change that Johns (2021) made to the semantic distinctiveness model architecture was modifying the representational assumptions of the model. Specifically, in Johns (2021), two types of representation types were tested within the model's framework: 1) word representations and 2) population representations. The word representation, initially introduced by Johns et al. (2020), involves a vector representing the count of how often each word appears within a specific contextual unit (either a discourse for DCD or a user for UCD). Thus, the word representation is fundamentally linguistic in nature; a context is represented by the words that occurred within that context.

In contrast, the population representation is based not on word usage patterns but on commenting patterns within a given context. For the DCD measure, the population context representation counts the number of comments each user made within a specific discourse, with the dimensionality of the representation being the number of users contained in the corpus. Consequently, the context representation for each discourse is a vector where each element represents a user, and the value of the element is the number of comments that the user made in that discourse (e.g., how many comments user X made in discourse Y).

For the UCD measure, the population context representation counts the number of comments each user made across various discourses. Each user's context representation is a vector where each element represents a specific discourse (e.g., r/AskReddit), and the value indicates how many comments the user made in that discourse. Because there are more users than discourses, the UCD measure receives more updates than the DCD measure.

https://doi.org/10.1017/S1366728925100229 Published online by Cambridge University Press

Population representation models were found to offer a significant advantage over their count-based counterparts and word representation models, suggesting that linguistic contexts encompass not only the words used but also communicative details, such as who produced the language and in which discourse. For concrete examples of how these representations are constructed, refer to Johns (2021) and Johns and Jones (2022). Importantly, the same advantages for the UCD and DCD measures that Johns (2021) found for young adult lexical retrieval data were found to generalize across the aging spectrum for monolingual lexical decision data (Johns et al., 2022).

### 1.4. Current study and hypotheses

Previous research has investigated SD in samples of young English speakers. Johns et al. (2016b) were the first to extend the semantic distinctiveness model to examine word recognition across aging and bilingualism. Differential language experience plays a key role in Gollan et al.'s (2008) frequency lag hypothesis and Ramscar et al.'s (2014) information accumulation perspective on aging. Therefore, exploring how CD may affect these groups is essential. Using a lexical decision task, they determined that bilinguals and older adults were more sensitive to SD information than younger monolinguals. They concluded that the unique language experiences of bilinguals and older adults lead to greater importance being placed on contextual information. In the present study, we extend these findings to examine the role of SD in picture naming ability across aging and bilingualism, and include L1 and L2 English speakers. Similar to previous work, CD measures are hypothesized to account for more variance across the groups than WF. Furthermore, we expect that the UCD-SD measure will outperform the DCD-SD measure.

Over time, older adults have increased their lexical knowledge, including contextual information (Ramscar et al., 2014). Their lexical organization may therefore rely more on this information source than it would for younger adults (Johns et al., 2016b). Similarly, bilinguals, who must decide which language(s) to use in a given situation (Gollan & Ferreira, 2009), develop a heightened ability to discriminate between contexts. This ability helps them organize their lexicon more effectively by using context to determine which language to speak. As a result, the variability of contexts plays a larger role in lexical organization (Johns et al., 2016b). In addition, because bilinguals must split their time between two languages, they have a lower level of experience with words in each language (Gollan et al., 2011). They may depend more on other linguistic information, such as contextual cues, for organizing their lexicon to compensate for their lower level of experience compared to monolinguals (Johns et al., 2016b).

Therefore, we hypothesized that bilinguals and older monolinguals would show greater sensitivity to the CD measure, compared to young monolinguals. Because older bilinguals have more language experience and acquired lexical information, we predicted that the SD measures would account for the most variance in this group.

Within the bilingual sample, we anticipated that L1 English speakers would be more sensitive to the CD measure because they have more English linguistic experience, including exposure to varying contexts. On the other hand, because L2 English speakers have less experience in English compared to L1 speakers, they rely on different strategies. We therefore, expected that the L2 English speakers would be more sensitive to WF and AoA than the CD measures.

## 2. Methods

### 2.1. Participants

The present study uses data from a larger project that aims to develop a new 30-item picture naming task that is appropriate for French, English and bilingual Canadians. Participants included the following groups: monolingual younger adults ($n$ = 49) and bilingual younger adults (further divided into L1 [$n$ = 22] and L2 English [$n$ = 21]), as well as monolingual older adults ($n$ = 64) and bilingual older adults (further divided into L1 [$n$ = 21] and L2 English ([$n$ = 31]). The sample size obtained in this study met the minimum required for conducting a Rasch analysis for the 30-item version of the naming task (see Linacre, 1994 for a detailed explanation). Younger participants were aged 18-30, while the older participants were aged 65 and above. Monolingual participants were fluent in English only, while bilingual participants were fluent in both English and French, and did not speak any other languages. Participants were categorized as fluent in a language if they reported their speaking, reading and auditory comprehension as at least 4, and writing as at least a 3 on a five-point Likert scale (1 = no ability at all; 2 = very little ability; 3 = moderate ability; 4 = very good ability; 5 = native-like ability). While self-ratings of language proficiency are commonly used in bilingualism research, there is evidence that these measures may not be completely reliable; bilinguals often underreport their proficiency (Tomoschuk et al., 2019; Wagner et al., 2022). In our experience, French speakers tend to rate their writing skills lower than English speakers do. Based on this pattern and previous work with this population, we accepted a rating of 3 ("moderate ability") as indicative of functional writing proficiency for group inclusion.

We first compared the performance of all bilinguals to that of monolinguals and then separated the bilingual sample into L1 and L2 English speakers. L1 English speakers acquired English first and French second, whereas L2 English speakers acquired French first and English second. All bilinguals had attained a high degree of proficiency in both languages before age 13. Proficiency data for the bilingual groups are presented in Table 1.

All participants were recruited through word of mouth and advertisements in community centers in Ottawa, Ontario, Canada. To determine eligibility, interested participants were contacted by telephone to discuss their language, education, and medical history. At the time of recruitment, participants self-reported no major neuropsychological problems and overall good health.

### 2.2. Materials

#### 2.2.1. Neuropsychological battery

To characterize participants' cognitive functioning across domains relevant to language processing and aging, each participant completed a neuropsychological assessment comprised of the following tests: Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), letter-number sequencing, a version of the Stroop color-word interference test (Stroop, 1935), forward and backward digit span subtest of the Wechsler Memory Scale (Wechsler, 1997), Wisconsin Card Sorting Task (WCST; Grant & Berg, 1948), Boston Naming Test (BNT; Kaplan et al., 1983), as well as category and letter verbal fluencies (Benton & Hamsher, 1976). This battery assessed general cognitive status, working memory, executive function and lexical access. These background measures were used to ensure that participants met inclusion criteria and to compare cognitive profiles across groups. Scores by participant groups are provided in Table 1.

https://doi.org/10.1017/S1366728925100229 Published online by Cambridge University Press

**Table 1.** Participants' demographic, neuropsychological and language characteristics (reported as mean ± standard deviation)

| | Younger adults | | | Older adults | | | |
|---|---|---|---|---|---|---|---|
| | | Bilingual | | | Bilingual | | |
| | Monolingual | L1 English | L2 English | Monolingual | L1 English | L2 English | Group comparisons* |
| Age (years) | 22.08 ± 1.94 | 21.90 ± 2.07 | 21.65 ± 2.30 | 72.19 ± 6.00 | 69.24 ± 4.87 | 73.03 ± 6.71 | Older adults > Young adults |
| Education (years) | 15.69 ± 1.21 | 15.59 ± 1.71 | 15.75 ± 1.33 | 15.20 ± 2.82 | 16.55 ± 2.35 | 16.17 ± 2.87 | NS |
| Sex (M/F) | 23/25 | 8/14 | 7/13 | 25/39 | 11/11 | 12/17 | NS |
| L2 AoA (years) | | 3.33 ± 2.25 | 5.05 ± 4.64 | | 4.42 ± 5.82 | 6.46 ± 4.85 | Older L2 > Young L1 |
| L1 listening proficiency (/5) | 4.96 ± 0.20 | 5.00 ± 0.00 | 5.00 ± 0.00 | 4.97 ± 0.18 | 4.86 ± 0.35 | 5.00 ± 0.00 | NS |
| L1 reading proficiency (/5) | 4.94 ± 0.32 | 4.90 ± 0.30 | 5.00 ± 0.00 | 4.95 ± 0.21 | 5.00 ± 0.00 | 4.86 ± 0.35 | NS |
| L1 speaking proficiency (/5) | 4.96 ± 0.20 | 4.95 ± 0.22 | 4.90 ± 0.31 | 4.98 ± 0.13 | 5.00 ± 0.00 | 4.90 ± 0.31 | NS |
| L1 writing proficiency (/5) | 4.96 ± 0.20 | 4.86 ± 0.36 | 4.80 ± 0.41 | 4.94 ± 0.30 | 5.00 ± 0.00 | 4.69 ± 0.54 | Older L2 < Older L1 |
| L2 listening proficiency (/5) | | 4.67 ± 0.58 | 4.75 ± 0.44 | | 4.77 ± 0.53 | 4.83 ± 0.38 | NS |
| L2 reading proficiency (/5) | | 4.43 ± 0.51 | 4.80 ± 0.52 | | 4.64 ± 0.49 | 4.86 ± 0.35 | Young L1 < Older L2 |
| L2 speaking proficiency (/5) | | 4.14 ± 0.66 | 4.70 ± 0.47 | | 4.64 ± 0.49 | 4.69 ± 0.47 | Young L1 < Young L2, Older L2 |
| L2 writing proficiency (/5) | | 3.95 ± 0.921 | 4.40 ± 0.681 | | 4.23 ± 0.752 | 4.76 ± 0.435 | Young L1, Older L1 < Older L2 |
| MoCA (/30) | 28.16 ± 1.41 | 28.14 ± 1.62 | 27.35 ± 1.35 | 27.63 ± 1.61 | 28.14 ± 1.55 | 27.28 ± 1.81 | NS |
| Stroop 1 | 102.72 ± 20.84 | 106.63 ± 16.93 | 110.30 ± 11.56 | 94.51 ± 15.02 | 98.64 ± 16.36 | 93.67 ± 14.24 | Younger L2 > Older L2, Older mono |
| Stroop 2 | 77.11 ± 15.01 | 74.55 ± 20.82 | 76.60 ± 7.97 | 65.79 ± 12.74 | 66.41 ± 12.13 | 58.00 ± 11.47 | Older adults < Young L2, Young Mono; Older L2 < Young L1 |
| Stroop 3 | 52.15 ± 11.88 | 51.00 ± 13.76 | 52.05 ± 7.19 | 34.30 ± 7.78 | 39.68 ± 8.63 | 34.55 ± 6.85 | Older adults < Young adults |
| Forward digit span (/16) | 11.04 ± 1.97 | 11.86 ± 2.03 | 10.20 ± 2.63 | 10.68 ± 1.89 | 10.86 ± 2.30 | 10.31 ± 1.80 | NS |
| Backward digit span (/14) | 6.84 ± 1.99 | 8.10 ± 2.39 | 7.55 ± 2.40 | 7.48 ± 2.20 | 8.09 ± 2.09 | 7.34 ± 2.56 | NS |
| BNT (/60) | 52.91 ± 3.52 | 46.20 ± 16.56 | 40.00 ± 15.74 | 54.00 ± 3.97 | 49.91 ± 11.91 | 49.25 ± 7.12 | Older L2 < Older mono; Young L2 < Young Mono |
| Verbal fluency total (FAS) | 41.67 ± 12.37 | 38.05 ± 10.40 | 34.70 ± 11.31 | 43.08 ± 10.84 | 41.82 ± 13.37 | 39.69 ± 12.59 | NS |
| Verbal fluency animals | 23.55 ± 5.84 | 25.24 ± 7.44 | 21.70 ± 6.51 | 21.29 ± 4.56 | 21.82 ± 5.74 | 18.28 ± 4.49 | Older L2 < Older mono, Young mono, Young L1 |
| WCST (/6) | 4.38 ± 0.98 | 4.47 ± 0.77 | 4.78 ± 0.43 | 3.60 ± 1.26 | 3.73 ± 1.03 | 3.63 ± 1.33 | Older mono < Young adults; Older L1, Older L2 < Young L1 |

*Note*: *All comparisons significant at $p < .05$. NS = not significant; Mono = monolingual, L1 = English first, L2 = English second, AoA = Age of Acquisition, MoCA = Montreal Cognitive Assessment, BNT = Boston Naming Test, WCST = Wisconsin Card Sorting Task.

**Montreal Cognitive Assessment**. The MoCA (Nasreddine et al., 2005) is a 10-minute cognitive screening test scored out of 30. It was designed to detect mild cognitive impairment and assess eight cognitive domains: short-term memory recall, visuospatial ability, executive function, attention, concentration, working memory, language, and orientation to time and place. A cut-off score of 24 was applied, in line with education-adjusted norms (i.e., with one additional point awarded for participants with <12 years of education) (Pugh et al., 2018).

**Letter-Number Sequencing**. The letter-number sequencing task is a measure of working memory and attention. In this task, participants heard a series of letters and numbers read aloud and were asked to reorder them by stating the numbers in ascending order, followed by the letters in alphabetical order (e.g., 4-O-8-H would be repeated as 4, 8, H, O). The task comprised seven items with three trials each, for a maximum score of 21. One point was awarded for each correct trial.

**Forward and Backward Digit Span**. Participants completed the forward and backward digit span subtests from the Wechsler Memory Scale (Wechsler, 1997). In the forward span, which assesses short-term memory, participants repeated 16 sequences of numbers in the order presented. In the backward span, which

https://doi.org/10.1017/S1366728925100229 Published online by Cambridge University Press

assesses working memory, they repeated 14 sequences in reverse order. One point was awarded for each correctly recalled sequence, with a maximum score of 16 for the forward span and 14 for the backward span.

**Stroop**. Participants completed the Stroop task (Stroop, 1935), which assesses inhibitory control, and is comprised of three conditions: word reading (Stroop 1), color naming (Stroop 2) and incongruent color naming (Stroop 3). Each condition consisted of a page displaying 120 stimuli. Participants were instructed to read the stimuli sequentially and were given 45 seconds per condition to name as many items as possible. The number of correct responses within the time limit was recorded for each condition.

**Wisconsin Card Sorting Task**. The WCST is a set-shifting test that assesses cognitive flexibility (Grant & Berg, 1948). Participants were given 64 cards, one at a time and were asked to organize each card according to three categories (color, shape and number). Participants were not told how to sort the cards; they were only informed whether their choice was "correct" or "incorrect" after they laid the card down. Cards were sorted first by color, then by shape and finally by number, with the category changing after 10 cards were correctly sorted. One point was awarded for each set of 10 consecutive correct responses, with a maximum score of six points.

**Boston Naming Test**. The BNT (Kaplan et al., 1983) comprises 60 line drawings displayed on a white background and arranged in increasing order of difficulty. Participants were asked to name each image while the researcher recorded their responses. There was no time limit for the task. Participants were scored out of 60.

**Verbal Fluencies**. Participants completed two fluencies: letter and category (Benton & Hamsher, 1976), which are measures of lexical access. The letter fluency task required participants to generate as many words as possible that begin with a given letter. Each participant completed this task for three letters: F, A, and S, for a final total letter fluency score. In the category fluency, participants were given a category (animals) and were asked to list as many words as they could that belonged to that category. Each fluency was recorded for 60 seconds, and one point was given for every acceptable word that the participant provided.

#### 2.2.2. 120-item naming task

The 120-item Naming Task includes high-quality digital images, 100 of which were selected from the colored Snodgrass and Vanderwart set (Rossion & Pourtois, 2001) and 20 of which were developed specifically for this task (see Appendix A, Table A1 for the full list of items and their lexical properties). The Snodgrass and Vanderwart images were chosen for their varying difficulty and strong name agreement, while the additional 20 images were created to match the Snodgrass and Vanderwart set but with increased naming difficulty. The images were displayed on a white background with Microsoft PowerPoint in the same randomized order for all participants. Participants were asked to name each image while the researcher logged their responses. If the participant provided the correct response, the researcher checked the "Correct Response" box. If the participant misinterpreted the picture, a stimulus cue was provided, and the participant could respond again (see Appendix A, Table A1 for the items and their corresponding cues). Stimulus cues were semantic; for example, the cue for the item "pomegranate" was that "it is a type of fruit." Any other responses provided by the participant were written down. One point was awarded for each correct, uncued response, for a maximum of 120 points. Of the 120 items, 17 images were removed from the subsequent analyses because the objects were multi-word, had low name agreement, or participants had difficulty visually identifying the item. The highest naming task score that participants were able to receive in this study was therefore 103.

### 2.3. Procedures

Participants completed two sessions at the Bruyère Health Research Institute in Ottawa, Ontario, Canada, that were each approximately 2 hours long. Over the two sessions, participants completed a neuropsychological battery, including the 120-item Naming Task. For monolinguals, language tasks were completed in English only, while bilingual participants completed all language tasks three times: in French, in English, and in a condition in which they could respond in either language. Ordering effects were eliminated by counterbalancing these language administrations. For bilingual participants, two of the language administrations were completed during the first session, while a third administration was completed during the second session. In the present analyses, the English-only scores were used for the bilingual participants. Participants received $10/hour compensation. The study procedures received ethical approval from the Research Ethics Board at the Bruyère Health Research Institute and the University of Ottawa. Prior to testing, the participants were briefed on the study procedures before providing written consent.

## 3. Results

A two-way ANOVA (see Table 2) examined the effect of age (younger and older) and language (monolingual and bilingual) on naming task scores. The main effects for age, $F(1, 209) = 9.82$, $p = .002$, $\eta^2 = .039$, and language, $F(1, 209) = 32.93$. $p < .001$, $\eta^2 = .130$, were both significant. However, there was no statistically significant interaction between the effects of age and language on naming task scores, $F(1, 209) = 0.704$, $p = 0.402$, $\eta^2 = .003$.

As a first pass at understanding the fit of the picture naming data and the various lexical variables (AoA, WF, DCD-SD and UCD-SD), Table 3 shows the correlations between the variables and behavior of the four participant groups. The table displays standard results, with the younger participant groups having higher correlations to the lexical strength variables than the older participant groups. Replicating the previously reported results of Johns (2021) and Johns et al. (2022), the CD measures had stronger correlations across all groups when compared to WF. Additionally, the UCD-SD measure outperformed the DCD-SD measure across all groups, also consistent with past results. AoA had the strongest correlation for the younger participant groups, while UCD-SD had the strongest correlation for the older participant groups. Since the UCD-SD variables provide a superior level of fit compared to the DCD-SD

**Table 2.** Mean scores and standard deviations across the different age groups for the English administration of the 103-item version of the naming task

| Age group | Mono versus Bil | Naming task /103 |
|---|---|---|
| Younger | *Mono* | 88.76 ± 6.12 |
| | *Bil* | 82.14 ± 9.95 |
| Older | *Mono* | 89.73 ± 5.65 |
| | *Bil* | 91.06 ± 4.65 |

**Table 3.** Correlations between naming performance for different participant groups and lexical variables

| Group | AoA | WF | DCD-SD | UCD-SD |
|---|---|---|---|---|
| Mono Younger | −.63 | .48 | .52 | .55 |
| Bil Younger | −.65 | .50 | .55 | .58 |
| Mono Older | −.38 | .35 | .39 | .40 |
| Bil Older | −.47 | .43 | .48 | .49 |

*Note*: Number of words = 103; all correlations significant at p < 0.001. AoA = Age of acquisition, WF = Word frequency, DCD-SD = Discourse contextual diversity modified by the semantic distinctiveness model, UCD-SD = User contextual diversity modified by the semantic distinctiveness model.

variable, UCD-SD is the only CD measure used in subsequent analyses.

To examine the separate effects of the different variables on the naming data, we conducted multiple sets of hierarchical linear regressions (HLRs), following the methodology of previous studies (Adelman et al., 2006; Adelman & Brown, 2008; Chang et al., 2023). Specifically, we conducted HLRs comparing WF and UCD-SD to determine which lexical strength variable accounts for the greatest amount of unique variance and to determine if the superiority of the UCD-SD measure over WF holds for picture naming data. An additional HLR was conducted comparing AoA and UCD-SD to determine which is the most powerful variable overall for the different age groups. In an HLR, the unique contribution of one variable over the competing variables was indexed by the percentage of $\Delta R^2$ in percent between the respective models (e.g., the percent increase in $\Delta R^2$ for WF and UCD-SD compared to just WF). See Chang and colleagues (2023) for more discussion of these analysis techniques.

The results of the regression analyses are contained in Figure 1. The top panel of this figure displays the amount of unique variance that WF and UCD-SD account for when compared against each other. Consistent with past results, it is found that UCD-SD accounts for the most unique variance while reducing the contribution of WF. This finding suggests that picture naming may operate under similar organizing principles as other lexical behaviors, such as lexical decision tasks (see also Van Assche et al., 2016). The bottom panel contrasts AoA and UCD-SD, which produced divergent results for the older and younger groups. Specifically, for the younger groups, AoA accounted for the most unique variance, while for the older groups, the UCD-SD measure accounted for greater amounts of unique variance for both monolinguals and bilinguals. This finding suggests that older adults' greater linguistic experience increases the sensitivity of the measures of lexical strength calculated by the UCD-SD measure.

However, as can be seen from Figure 1, there are not many divergences in fit for the bilingual groups compared to the monolingual groups. To gain a better understanding of the performance of the bilingual participant groups, the young and older bilingual groups were split into two groups depending on whether they were L1 English speakers or L2 English speakers (see Table 4). A two-way ANOVA examined the effect of age (younger and older) and language dominance of the bilingual participants (L1 English and L2 English) on naming task scores. A similar pattern was observed in these split groups, where main effects of age, $F(1, 89) = 7.41$, $p = .008$, $\eta^2 = .064$, and language dominance $F(1, 89) = 18.54$, $p < .001$, $\eta^2 = .161$, were both significant. Again, there was no statistically significant interaction between the effects of age and

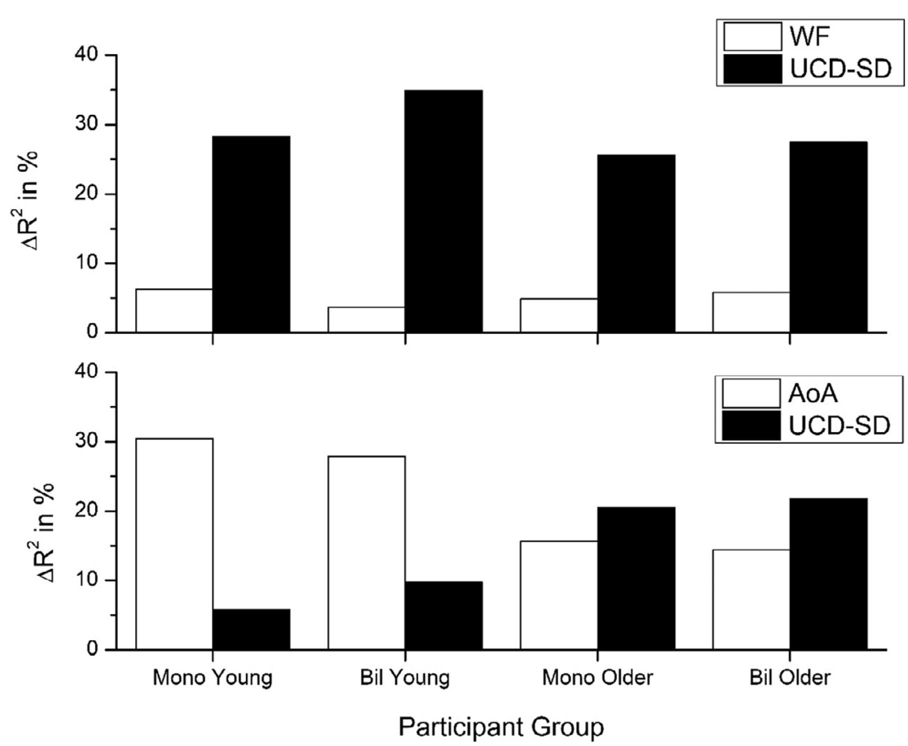


**Figure 1.** Hierarchical linear regression analyses comparing the amount of unique variance accounted for when comparing WF and UCD-SD (top panel) and AoA and UCD-SD (bottom panel) across the different age groups.

**Table 4.** Mean scores and standard deviations across the split bilingual groups for the English administration of the 103-item version of the naming task

| Age group | L1 versus L2 English bilingual | Naming task /103 |
|---|---|---|
| Younger | *L1 English* | 86.14 ± 8.51 |
| | *L2 English* | 76.80 ± 11.34 |
| Older | *L1 English* | 89.73 ± 5.65 |
| | *L2 English* | 83.41 ± 6.99 |

language dominance on naming task scores, $F(1, 89) = 0.443$, $p = 0.507$, $\eta^2 = .004$.

Hierarchical regression analyses with the same comparisons as contained in Figure 1 were used to assess the amount of unique variance that the different lexical variables account for across the split bilingual groups. The results of the regression on the split bilingual groups are shown in Figure 2, with the top panel displaying the WF versus UCD-SD comparison and the bottom panel displaying the AoA versus UCD-SD comparison. For the WF and UCD-SD comparison, the results are similar to those shown in Figure 1, with UCD-SD accounting for the most unique variance across all groups. However, there is a difference in the AoA versus UCD-SD comparison, where for the older L1 English group, the UCD-SD variable accounts for more variance than AoA, while for the older L2 English group, AoA accounts for the most unique variance, similar to the young participants. This finding suggests that the older L2 English group's more limited lexical experience with English reduces the contribution of the lexical strength values of the UCD-SD measure.

## 4. Discussion

The present study aimed to examine the role of CD in picture naming in aging and bilingualism. Picture naming ability was measured in a sample of young and older English monolinguals and English–French bilinguals. A comparison between WF, AoA, DCD and UCD was conducted to determine which information source best fits the lexical organization of the groups. Overall, these findings indicate the necessity of recognizing how linguistic experiences are shaped by age and bilingualism. As previously mentioned, Ramscar et al. (2014) and Gollan et al. (2008) emphasize the importance of including contextual information when understanding differences in lexical access.

Akin to Johns (2021), it was determined that the UCD-SD offers a superior level of fit compared to the DCD-SD across all groups, and was especially superior to the classic WF variable. These results replicate previous findings that suggest that the CD effect is stronger than the WF effect (Adelman et al., 2006; Adelman & Brown, 2008; Brysbaert & New, 2009; Johns et al., 2016b, 2022; Vergara-Martínez et al., 2017). The UCD measure is based on how likely a large group of people is to use particular words and therefore provides a more social type of information than the DCD measure (Johns, 2021). Usage-based theories of language acquisition such as Tomasello (2003) propose that language is learned through observation and understanding how others use language in a communicative environment (Johns et al., 2022).

As anticipated, UCD-SD accounted for the most variance in all groups compared to WF, as has been seen consistently in the literature (e.g., Adelman et al., 2006; Adelman & Brown, 2008; Brysbaert & New, 2009; Johns et al., 2016b, 2022). This finding indicates that picture naming and other lexical behaviors may share common organizing principles. Larger contextual measures of

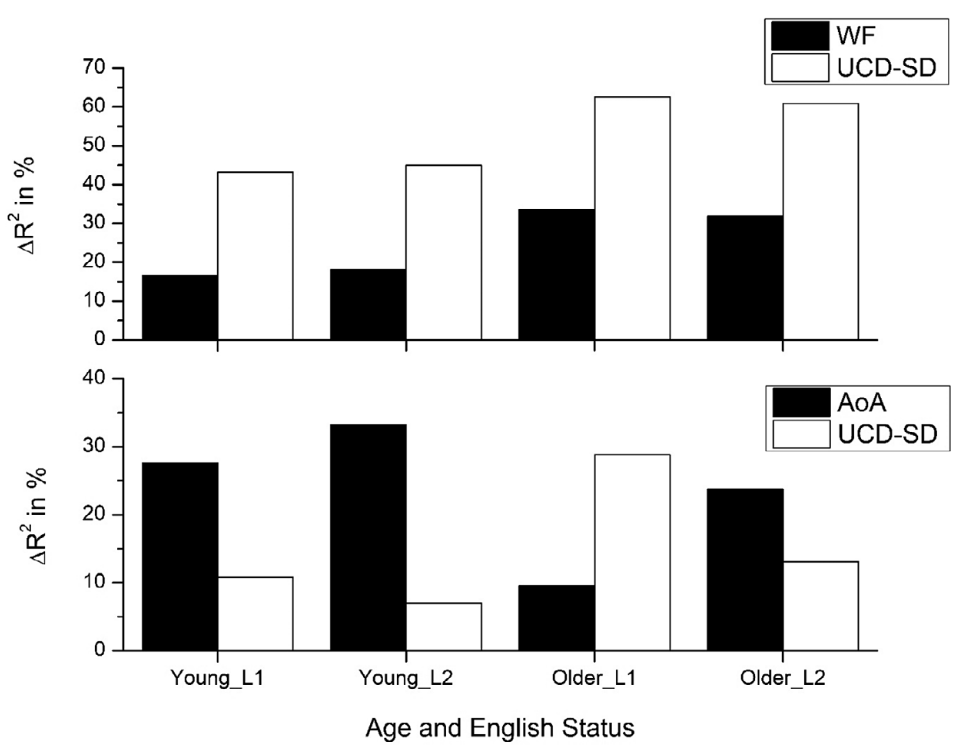


**Figure 2.** Hierarchical linear regression analyses comparing the amount of unique variance accounted for when comparing WF and UCD-SD (top panel) and AoA and UCD-SD (bottom panel) for the split bilingual groups.

word usage can provide a more precise understanding of whether a word has been encountered previously (Johns, 2021). In addition, these findings contradict the suggestion that CD could substitute for WF in models of lexical organization and word integration (Plummer et al., 2014). As Vergara-Martínez and colleagues (2017) observed, CD and WF effects may be correlated, but are separate. Therefore, we suggest that people may use more abstract cues, like SD and CD, to organize their lexicon, compared to repetition-based learning mechanisms like WF (Johns, 2021; Senaldi et al., 2022). These results support a theoretical shift toward models of lexical access that emphasize socially grounded, semantically enriched, and context-sensitive measures.

Interestingly, we found that the strength of the CD measure was affected by age. Ramscar et al. (2014) proposed that older adults have more accumulated linguistic knowledge compared to younger adults. Our findings support this notion: older adults were more sensitive to CD information than younger adults, who instead relied more on AoA. This result indicates that strategies used for lexical encoding shift over the lifespan, with older adults' additional lexical experience resulting in a greater sensitivity to CD. Using a lexical decision paradigm, Johns et al. (2016b) similarly demonstrated that older adults used information derived from the semantic distinctiveness model to a greater extent than younger adults. Qiu and Johns (2020) propose that older adults' increased language experience allows more SD information to be encoded in the lexical memory compared to younger adults. Regardless, these findings emphasize the importance of semantic information, whether that be in the form of CD or AoA, over WF in models of lexical organization for both younger and older adults.

Because the bilingual groups completed the naming task three times, it is likely that the WF effect (or at least an immediate version of this effect) would be stronger for the bilinguals than the monolinguals. However, our results show no significant differences in WF across the monolingual and bilingual groups, providing further evidence to support the importance of semantic and contextual information in lexical encoding for bilinguals over WF.

To gain a deeper understanding of the impact of these variables on picture naming performance in bilinguals, we divided the bilingual sample into L1 and L2 English speakers. Because L1 English speakers have more linguistic experience in English, we expected them to be more sensitive to the CD measure than L2 English speakers, who we hypothesized would rely more on other information such as WF and AoA. This rationale held for the AoA measure, with L1 speakers relying more on SD and CD information than L2 speakers, while the L2 speakers relied more on AoA.

When the effects of AoA and WF were compared to those of UCD-SD, contrasting results were found. L2 speakers relied more on AoA than UCD-SD information, likely because they have less lexical experience in English, which was the language of testing. Among the four groups, only older L1 English speakers relied more on CD information than on AoA. This may be because the older L1 English speakers, being both bilingual and older, have accumulated more semantic knowledge over time than other groups. However, no such effect was found when comparing WF and UCD-SD. Rather, all groups were more sensitive to the UCD-SD measure than to WF. That is, even when bilinguals are tested in their L2, they are more sensitive to contextual than WF information. This division into L1 and L2 bilinguals offers a novel contribution to the bilingualism literature, showing that the type and extent of English language experience modulate sensitivity to different lexical predictors.

Finally, it is important to note that the WF measure consistently provided the worst fit for all the groups in this study. Research on WF and picture naming ability has shown mixed results, with some studies highlighting its importance (e.g., Alario et al., 2004; Barry et al., 1997; Bonin et al., 2003; Cuetos et al., 1999; Cuitiño et al., 2019; Ellis & Morrison, 1998; Snodgrass & Yuditsky, 1996), while others have failed to find any effect (e.g., Bonin et al., 2002; Dell'acqua et al., 2000; Valente et al., 2014). Our findings replicate those reported in the latter studies, suggesting that frequency measures may not be as useful or informative as initially believed and that more social aspects of language acquisition, like CD, play a larger role in picture naming. Beyond this, these findings emphasize the importance of semantic information, which encompasses both AoA and CD, in the organization of the lexicon over WF.

One limitation in the present study is that only accuracy data were available; including naming latency data would shed further light on the phenomena examined here. Other accuracy-based measures, such as error patterns and cued responses, should also be investigated, as these data could provide insight into how CD impacts word retrieval. However, in the present study, only 0.28% of all naming task responses were cued, limiting their analytical value. Studies investigating picture naming in other populations, such as individuals with aphasia, where cued responses are more common (Meteyard & Bose, 2018), may be better suited to explore this aspect. Furthermore, this study investigated a limited number of psycholinguistic variables: UCD, DCD, AoA and WF. Other variables, such as name agreement (Alario et al., 2004; Barry et al., 1997; Bonin et al., 2002, 2003; Cuetos et al., 1999; Cuitiño et al., 2019; Dell'acqua et al., 2000; Ellis & Morrison, 1998; Khwaileh et al., 2018; Perret & Bonin, 2019; Snodgrass & Yuditsky, 1996; Valente et al., 2014), imageability (Alario et al., 2004; Ballot et al., 2021; Bonin et al., 2002, 2003; Ellis & Morrison, 1998; Khwaileh et al., 2018; Perret & Bonin, 2019) and concept familiarity (Cuetos et al., 1999; Cuitiño et al., 2019; Ellis & Morrison, 1998; Khwaileh et al., 2018; Perret & Bonin, 2019; Snodgrass & Yuditsky, 1996), have been shown to impact picture naming ability as well. Finally, future research should also examine more fine-grained aspects of language usage in bilingual speakers, such as code-switching behavior and variation in the languages a person uses in different settings (e.g., at home, at work, with friends). This variation can be computed as language entropy and is based on the proportion of time a person uses different languages (Gullifer et al., 2021). Exploring the interconnectedness of CD and language entropy may also be relevant, as both contribute to a bilingual's language experience.

## 5. Conclusion

Previous research has shown that context in lexical organization plays an important role in many aspects of lexical processing (Adelman et al., 2006; Chang, Jones & Johns, in press; Jones et al., 2017; 2012). In an initial inquiry into the relationship between context, bilingualism and aging, Johns et al. (2016b) found that bilinguals and older adults were more sensitive to SD information during a lexical decision task than monolinguals and younger adults. The present study extends these findings by demonstrating that the same effects were present for picture naming ability. We determined that older adults and L1 English speakers are more sensitive to CD information than younger adults and L2 English speakers, and that the latter group instead relies primarily on AoA. These findings emphasize how individual linguistic experiences, which are shaped by both age and bilingualism, impact lexical

https://doi.org/10.1017/S1366728925100229 Published online by Cambridge University Press

processing. Moreover, the present study illustrates the theoretical importance of delving deeper into the language function of older adults and bilinguals.

**Data availability statement.** Please contact the corresponding author to request access to the de-identified data used in this study.

**Acknowledgements.** This research was supported by an Alzheimer Society of Canada Research Grant awarded to Vanessa Taler (Grant #1423).

**Competing interest.** The authors declare none.

## References

**Adelman, J. S.**, & **Brown, G. D. A.** (2008). Modeling lexical decision: The form of frequency and diversity effects. *Psychological Review*, **115**(1), 214–227. https://doi.org/10.1037/0033-295X.115.1.214.

**Adelman, J. S.**, **Brown, G. D. A.**, & **Quesada, J. F.** (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science*, **17**(9), 814–823. https://doi.org/10.1111/j.1467-9280.2006.01787.x.

**Alario, F.-X.**, **Ferrand, L.**, **Laganaro, M.**, **New, B.**, **Frauenfelder, U. H.**, & **Segui, J.** (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, **36**(1), 140–155. https://doi.org/10.3758/BF03195559.

**Anderson, J. R.** (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, **6**(4), 451–474. https://doi.org/10.1016/0010-0285(74)90021-8.

**Anderson, J. R.**, & **Milson, R.** (1989). Human memory: An adaptive perspective. *Psychological Review*, **96**(4), 703–719. https://doi.org/10.1037/0033-295X.96.4.703.

**Anderson, J. R.**, & **Schooler, L. J.** (1991). Reflections of the environment in memory. *Psychological Science*, **2**(6), 396–408. https://doi.org/10.1111/j.1467-9280.1991.tb00174.x.

**Ardila, A.**, & **Rosselli, M.** (1989). Neuropsychological characteristics of normal aging. *Developmental Neuropsychology*, **5**(4), 307–320. https://doi.org/10.1080/87565648909540441.

**Ballot, C.**, **Mathey, S.**, & **Robert, C.** (2021). Age-related evaluations of imageability and subjective frequency for 1286 neutral and emotional French words: Ratings by young, middle-aged, and older adults. *Behavior Research Methods*, **54**(1), 196–215. https://doi.org/10.3758/s13428-021-01621-6.

**Balota, D. A.**, **Yap, M. J.**, **Hutchison, K. A.**, **Cortese, M. J.**, **Kessler, B.**, **Loftis, B.**, **Neely, J. H.**, **Nelson, D. L.**, **Simpson, G. B.**, & **Treiman, R.** (2007). The English lexicon project. *Behavior Research Methods*, **39**(3), 445–459. https://doi.org/10.3758/BF03193014.

**Barry, C.**, **Morrison, C. M.**, & **Ellis, A. W.** (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology*, **50**(3), 560–585. https://doi.org/10.1080/783663595.

**Benton, A. L.** & **Hamsher, K.** (1976). *Multilingual aphasia examination manual*. University of Iowa, Iowa City: AJA Associates.

**Bialystok, E.**, **Craik, F.**, & **Luk, G.** (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **34**(4), 859–873. https://doi.org/10.1037/0278-7393.34.4.859.

**Bonin, P.**, **Chalard, M.**, **Méot, A.**, & **Fayol, M.** (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, **93**(1), 89–114. https://doi.org/10.1348/000712602162463.

**Bonin, P.**, **Peereman, R.**, **Malardier, N.**, **Méot, A.**, & **Chalard, M.** (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, **35**(1), 158–167. https://doi.org/10.3758/BF03195507.

**Brysbaert, M.**, & **Ellis, A. W.** (2016). Aphasia and age of acquisition: Are early-learned words more resilient? *Aphasiology*, **30**(11), 1240–1263. https://doi.org/10.1080/02687038.2015.1106439.

**Brysbaert, M.**, & **Ghyselinck, M.** (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition*, **13**(7–8), 992–1011. https://doi.org/10.1080/13506280544000165.

**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. https://doi.org/10.1177/0963721417727521.

**Brysbaert, M.**, & **New, B.** (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, **41**(4), 977–990. https://doi.org/10.3758/BRM.41.4.977.

**Brysbaert, M.**, **Warriner, A. B.**, & **Kuperman, V.** (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, **46**(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5.

**Burke, D. M.**, & **Mackay, D. G.** (1997). Memory, language, and ageing. *Philosophical Transactions: Biological Sciences*, **352**(1363), 1845–1856.

**Burke, D. M.**, & **Shafto, M. A.** (2004). Aging and language production. *Current Directions in Psychological Science*, **13**(1), 21–24. https://doi.org/10.1111/j.0963-7214.2004.01301006.x.

**Chang, M.**, **Jones, M. N.**, & **Johns, B. T.** (2023). Comparing word frequency, semantic diversity, and semantic distinctiveness in lexical organization. *Journal of Experimental Psychology: General*, **152**(6), 1814–1823. https://doi.org/10.1037/xge0001407.

**Christ, S. E.**, **White, D. A.**, **Mandernach, T.**, & **Keys, B. A.** (2001). Inhibitory control across the life span. *Developmental Neuropsychology*, **20**(3), 653–669. https://doi.org/10.1207/S15326942DN2003_7.

**Coltheart, M.**, **Rastle, K.**, **Perry, C.**, **Langdon, R.**, & **Ziegler, J.** (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, **108**(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204.

**Corps, R. E.**, & **Meyer, A. S.** (2023). Word frequency has similar effects in picture naming and gender decision: A failure to replicate Jescheniak and Levelt (1994). *Acta Psychologica*, **241**, 104073. https://doi.org/10.1016/j.actpsy.2023.104073.

**Costa, A.**, & **Caramazza, A.** (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition*, **2**(3), 231–244. https://doi.org/10.1017/S1366728999000334.

**Cuetos, F.**, **Ellis, A. W.**, & **Alvarez, B.** (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, **31**(4), 650–658. https://doi.org/10.3758/BF03200741.

**Cuitiño, M. M.**, **Soriano, F. G.**, **Jaichenco, V.**, **Steeb, B.**, & **Barreyro, J. P.** (2019). Predictors of picture naming and picture categorization in Spanish. *East European Journal of Psycholinguistics*, **6**(1), 6–18. https://doi.org/10.29038/eejpl.2019.6.1.cui.

**Dell'acqua, R.**, **Lotto, L.**, & **Job, R.** (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, **32**(4), 588–615. https://doi.org/10.3758/BF03200832.

**Ellis, A. W.**, & **Morrison, C. M.** (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, **24**(2), 515–523. https://doi.org/10.1037/0278-7393.24.2.515.

**Feyereisen, P.** (1997). A meta-analytic procedure shows an age-related decline in picture naming: Comments on Goulet, Ska, and Kahn (1994). *Journal of Speech, Language & Hearing Research*, **40**(6), 1328–1333. https://doi.org/10.1044/jslhr.4006.1328.

**Goldinger, S. D.** (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, **105**(2), 251–279. https://doi.org/10.1037/0033-295X.105.2.251.

**Gollan, T. H.**, **Fennema-Notestine, C.**, **Montoya, R. I.**, & **Jernigan, T. L.** (2007). The bilingual effect on Boston naming test performance. *Journal of the International Neuropsychological Society*, **13**(2), 197–208. https://doi.org/10.1017/S1355617707070038.

**Gollan, T. H.**, & **Ferreira, V. S.** (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **35**(3), 640–665. https://doi.org/10.1037/a0014981.

**Gollan, T. H.**, **Montoya, R. I.**, & **Bonanni, M. P.** (2005a). Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology*, **19**(3), 278–287. https://doi.org/10.1037/0894-4105.19.3.278.

**Gollan, T. H.**, **Montoya, R. I.**, **Cera, C.**, & **Sandoval, T. C.** (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, **58**(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001.

**Gollan, T. H.**, **Montoya, R. I.**, **Fennema-Notestine, C.**, & **Morris, S. K.** (2005b). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, **33**(7), 1220–1234. https://doi.org/10.3758/BF03193224.

**Gollan, T. H.**, **Slattery, T. J.**, **Goldenberg, D.**, **Van Assche, E.**, **Duyck, W.**, & **Rayner, K.** (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, **140**(2), 186–209. https://doi.org/10.1037/a0022256.

**Grant, D. A.**, & **Berg, E.** (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, **38**(4), 404–411. https://doi.org/10.1037/h0059831.

**Green, D. W.** (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, **1**(2), 67–81. https://doi.org/10.1017/S1366728998000133.

**Gullifer, J. W.**, **Kousaie, S.**, **Gilbert, A. C.**, **Grant, A.**, **Giroud, N.**, **Coulter, K.**, **Klein, D.**, **Baum, S.**, **Phillips, N.**, & **Titone, D.** (2021). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, **42**(2), 245–278. https://doi.org/10.1017/S0142716420000521.

**Hermans, D.**, **Bongaerts, T.**, **De Bot, K.**, & **Schreuder, R.** (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, **1**(3), 213–229. https://doi.org/10.1017/S1366728998000364.

**Hernandez, A.**, & **Li, P.** (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, **133**, 638–650. https://doi.org/10.1037/0033-2909.133.4.638.

**Ivanova, I.**, & **Costa, A.** (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, **127**(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003.

**Ivnik, R. J.**, **Smith, G. E.**, **Malec, J. F.**, **Petersen, R. C.**, & **Tangalos, E. G.** (1995). Long-term stability and intercorrelations of cognitive abilities in older persons. *Psychological Assessment*, **7**(2), 155–161. https://doi.org/10.1037/1040-3590.7.2.155.

**Jescheniak, J. D.**, & **Levelt, W. J. M.** (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, **20**(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824.

**Johns, B. T.** (2021). Disentangling contextual diversity: Communicative need as a lexical organizer. *Psychological Review*, **128**(3), 525–557. https://doi.org/10.1037/rev0000265.

**Johns, B. T.**, **Dye, M.**, & **Jones, M. N.** (2016a). The influence of contextual diversity on word learning. *Psychonomic Bulletin & Review*, **23**(4), 1214–1220. https://doi.org/10.3758/s13423-015-0980-7.

**Johns, B. T.**, **Dye, M.**, & **Jones, M. N.** (2020). Estimating the prevalence and diversity of words in written language. *Quarterly Journal of Experimental Psychology*, **73**(6), 841–855. https://doi.org/10.1177/1747021819897560.

**Johns, B. T.**, & **Jones, M. N.** (2008). Predicting word-naming and lexical decision times from a semantic space model. *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, **30**, 279–285.

**Johns, B. T.**, & **Jones, M. N.** (2022). Content matters: Measures of contextual diversity must consider semantic content. *Journal of Memory and Language*, **123**, 104313. https://doi.org/10.1016/j.jml.2021.104313.

**Johns, B. T.**, **Sheppard, C. L.**, **Jones, M. N.**, & **Taler, V.** (2016b). The role of semantic diversity in word recognition across aging and bilingualism. *Frontiers in Psychology*, **7**, 1–11. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00703.

**Johns, B. T.**, **Taler, V.**, & **Jones, M. N.** (2022). Contextual dynamics in lexical encoding across the ageing spectrum: A simulation study. *Quarterly Journal of Experimental Psychology*, **76**(9), 2164–2182. https://doi.org/10.1177/17470218221145685.

**Jones, M. N.**, **Dye, M.**, & **Johns, B. T.** (2017). Context as an organizing principle of the lexicon. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 239–283). Elsevier. https://doi.org/10.1016/bs.plm.2017.03.008.

**Jones, M. N.**, **Johns, B. T.**, & **Recchia, G.** (2012). The role of semantic diversity in lexical organization. *Canadian Journal of Experimental Psychology*, **66**(2), 115–124. https://doi.org/10.1037/a0026727.

**Kaplan, E.**, **Goodglass, H.**, & **Weintraub, S.** (1983). *Boston naming test*. Lea & Febiger.

**Kavé, G.**, **Kukulansky-Segal, D.**, **Avraham, A.**, **Herzberg, O.**, & **Landa, J.** (2010). Searching for the right word: Performance on four word-retrieval tasks across childhood. *Child Neuropsychology*, **16**(6), 549–563. https://doi.org/10.1080/09297049.2010.485124.

**Khwaileh, T.**, **Mustafawi, E.**, **Herbert, R.**, & **Howard, D.** (2018). Gulf Arabic nouns and verbs: A standardized set of 319 object pictures and 141 action pictures, with predictors of naming latencies. *Behavior Research Methods*, **50**(6), 2408–2425. https://doi.org/10.3758/s13428-018-1019-6.

**Kohnert, K. J.**, **Hernandez, A. E.**, & **Bates, E.** (1998). Bilingual performance on the Boston naming test: Preliminary norms in Spanish and English. *Brain and Language*, **65**(3), 422–440. https://doi.org/10.1006/brln.1998.2001.

**Kroll, J. F.**, **Bobb, S. C.**, **Misra, M.**, & **Guo, T.** (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, **128**(3), 416–430. https://doi.org/10.1016/j.actpsy.2008.02.001.

**Kumar, A. A.** (2021). Semantic memory: A review of methods, models, and current challenges. *Psychonomic Bulletin & Review*, **28**(1), 40–80. https://doi.org/10.3758/s13423-020-01792-x.

**Levelt, W. J. M.**, **Roelofs, A.**, & **Meyer, A. S.** (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, **22**(1), 1–38. https://doi.org/10.1017/S0140525X99001776.

**Levy, B. J.**, **McVeigh, N. D.**, **Marful, A.**, & **Anderson, M. C.** (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science*, **18**(1), 29–34. https://doi.org/10.1111/j.1467-9280.2007.01844.x.

**Linacre, J.** (1994). Sample size and item calibration stability. *Rasch Management Transactions*, **7**(8), 328. https://www.rasch.org/rmt/rmt74m.htm

**Linck, J. A.**, **Kroll, J. F.**, & **Sunderman, G.** (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, **20**(12), 1507–1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x.

**Lund, K.**, & **Burgess, C.** (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, **28**(2), 203–208. https://doi.org/10.3758/BF03204766.

**Mackay, A. J.**, **Connor, L. T.**, **Albert, M. L.**, & **Obler, L. K.** (2002). Noun and verb retrieval in healthy aging. *Journal of the International Neuropsychological Society*, **8**(6), 764–770. https://doi.org/10.1017/S1355617702860040.

**Mägiste, E.** (1979). The competing language systems of the multilingual: A developmental study of decoding and encoding processes. *Journal of Verbal Learning and Verbal Behavior*, **18**(1), 79–89. https://doi.org/10.1016/S0022-5371(79)90584-X.

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

**Misra, M.**, **Guo, T.**, **Bobb, S. C.**, & **Kroll, J. F.** (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, **67**(1), 224–237. https://doi.org/10.1016/j.jml.2012.05.001.

**Mitrushina, M.**, & **Satz, P.** (1995). Repeated testing of normal elderly with the Boston naming test. *Aging Clinical and Experimental Research*, **7**(2), 123–127. https://doi.org/10.1007/BF03324301.

**Morton, J.** (1969). Interaction of information in word recognition. *Psychological Review*, **76**(2), 165–178. https://doi.org/10.1037/h0027366.

**Murray, W. S.**, & **Forster, K. I.** (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review*, **111**(3), 721–756. https://doi.org/10.1037/0033-295X.111.3.721.

**Nasreddine, Z. S.**, **Phillips, N. A.**, **Bédirian, V.**, **Charbonneau, S.**, **Whitehead, V.**, **Collin, I.**, **Cummings, J. L.**, & **Chertkow, H.** (2005). The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, **53**(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x.

**Navarrete, E.**, **Basagni, B.**, **Alario, F.-X.**, & **Costa, A.** (2006). Does word frequency affect lexical selection in speech production? *Quarterly Journal of Experimental Psychology*, **59**(10), 1681–1690. https://doi.org/10.1080/17470210600750558.

**Nicholas, L. E.**, **Brookshire, R. H.**, **Maclennan, D. L.**, **Schumacher, J. G.**, & **Porrazzo, S. A.** (1989). Revised administration and scoring procedures for the Boston naming test and norms for non-brain-damaged adults. *Aphasiology*, **3**(6), 569–580. https://doi.org/10.1080/02687038908249023.

**Norris, D.** (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, **113**(2), 327–357. https://doi.org/10.1037/0033-295X.113.2.327.

**Paesen, L.**, & **Leijten, M.** (2019). Name agreement and naming latencies for typed picture naming in aging adults. *Clinical Linguistics & Phonetics*, **33**(10–11), 930–948. https://doi.org/10.1080/02699206.2019.1590734.

**Perret, C.**, & **Bonin, P.** (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, **51**(6), 2533–2545. https://doi.org/10.3758/s13428-018-1100-1.

**Philipp, A. M.**, **Gade, M.**, & **Koch, I.** (2007). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology*, **19**(3), 395–416. https://doi.org/10.1080/09541440600758812.

**Plummer, P.**, **Perea, M.**, & **Rayner, K.** (2014). The influence of contextual diversity on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **40**(1), 275–283. https://doi.org/10.1037/a0034058.

**Pugh, E. A.**, **Kemp, E. C.**, **van Dyck, C. H.**, **Mecca, A. P.**, & **Sharp, E. S.** (2018). Effects of normative adjustments to the Montreal cognitive assessment. *The American Journal of Geriatric Psychiatry*, **26**(12), 1258–1267. https://doi.org/10.1016/j.jagp.2018.09.009.

**Qiu, M.**, & **Johns, B. T.** (2020). Semantic diversity in paired-associate learning: Further evidence for the information accumulation perspective of cognitive aging. *Psychonomic Bulletin & Review*, **27**(1), 114–121. https://doi.org/10.3758/s13423-019-01691-w.

**Rabovsky, M.**, **Sommer, W.**, & **Abdel Rahman, R.** (2012). The time course of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, **6**, 11. https://doi.org/10.3389/fnhum.2012.00011.

**Ramscar, M.**, **Hendrix, P.**, **Shaoul, C.**, **Milin, P.**, & **Baayen, H.** (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, **6**(1), 5–42. https://doi.org/10.1111/tops.12078.

**Roberts, P. M.**, **Garcia, L. J.**, **Desrochers, A.**, & **Hernandez, D.** (2002). English performance of proficient bilingual adults on the Boston naming test. *Aphasiology*, **16**(4–6), 635–645. https://doi.org/10.1080/02687030244000220.

**Roelofs, A.** (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, **42**(1), 107–142. https://doi.org/10.1016/0010-0277(92)90041-F.

**Rosselli, M.**, **Ardila, A.**, & **Rosas, P.** (1990). Neuropsychological assessment in illiterates: II. Language and praxic abilities. *Brain and Cognition*, **12**(2), 281–296. https://doi.org/10.1016/0278-2626(90)90020-O.

**Rossion, B.**, & **Pourtois, G.** (2001). Revisiting snodgrass and Vanderwart's object database: Color and texture improve object recognition. *Journal of Vision*, **1**(3), 413. https://doi.org/10.1167/1.3.413.

**Senaldi, M. S. G.**, **Titone, D. A.**, & **Johns, B. T.** (2022). Determining the importance of frequency and contextual diversity in the lexical organization of multiword expressions. *Canadian Journal of Experimental Psychology*, **76**(2), 87–98. https://doi.org/10.1037/cep0000271.

**Shafto, M. A.**, **Burke, D. M.**, **Stamatakis, E. A.**, **Tam, P. P.**, & **Tyler, L. K.** (2007). On the tip-of-the-tongue: Neural correlates of increased word-finding failures in normal aging. *Journal of Cognitive Neuroscience*, **19**(12), 2060–2070. https://doi.org/10.1162/jocn.2007.19.12.2060.

**Shafto, M. A.**, & **Tyler, L. K.** (2014). Language in the aging brain: The network dynamics of cognitive decline and preservation. *Science*, **346**(6209), 583–587. https://doi.org/10.1126/science.1254404.

**Shaoul, C.**, & **Westbury, C.** (2010). Exploring lexical co-occurrence space using HiDEx. *Behavior Research Methods*, **42**(2), 393–413. https://doi.org/10.3758/BRM.42.2.393.

**Sheppard, C.**, **Kousaie, S.**, **Monetta, L.**, & **Taler, V.** (2016). Performance on the Boston naming test in bilinguals. *Journal of the International Neuropsychological Society*, **22**(3), 350–363. https://doi.org/10.1017/S135561771500123X.

**Silagi, M. L.**, **Bertolucci, P. H. F.**, & **Ortiz, K. Z.** (2015). Naming ability in patients with mild to moderate Alzheimer's disease: What changes occur with the evolution of the disease? *Clinics*, **70**(6), 423–428. https://doi.org/10.6061/clinics/2015(06)07.

**Snodgrass**, & **Yuditsky, T.** (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments & Computers*, **28**(4), 516–536. https://doi.org/10.3758/BF03200540.

**Stroop, J. R.** (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**(6), 643–662. https://doi.org/10.1037/h0054651.

**Tomasello, M.** (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

**Tomoschuk, B.**, **Ferreira, V. S.**, & **Gollan, T. H.** (2019). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, **22**(3), 516–536. https://doi.org/10.1017/S1366728918000421.

**Valente, A.**, **Bürki, A.**, & **Laganaro, M.** (2014). ERP correlates of word production predictors in picture naming: A trial by trial multiple regression analysis from stimulus onset to response. *Frontiers in Neuroscience*, **8**, 1–13. https://doi.org/10.3389/fnins.2014.00390.

**Van Assche, E.**, **Duyck, W.**, & **Gollan, T. H.** (2016). Linking recognition and production: Cross-modal transfer effects between picture naming and lexical decision during first and second language processing in bilinguals. *Journal of Memory and Language*, **89**, 37–54. https://doi.org/10.1016/j.jml.2016.02.003.

**Vergara-Martínez, M.**, **Comesaña, M.**, & **Perea, M.** (2017). The ERP signature of the contextual diversity effect in visual word recognition. *Cognitive, Affective, & Behavioral Neuroscience*, **17**(3), 461–474. https://doi.org/10.3758/s13415-016-0491-7.

**Wagner, D.**, **Bialystok, E.**, & **Grundy, J. G.** (2022). What is a language? Who is bilingual? Perceptions underlying self-assessment in studies of bilingualism. *Frontiers in Psychology*, **13**, 863991. https://doi.org/10.3389/fpsyg.2022.863991.

**Wang, J.**, & **Chen, B.** (2020). A database of Chinese-English bilingual speakers: Ratings of the age of acquisition and familiarity. *Frontiers in Psychology*, **11**, 554785. https://doi.org/10.3389/fpsyg.2020.554785.

**Wang, J.**, **Jiang, X.**, & **Chen, B.** (2023). Second language age of acquisition effects in a word naming task: A regression analysis of ERP data. *Journal of Neurolinguistics*, **66**, 1–11. https://doi.org/10.1016/j.jneuroling.2023.101125

**Wechsler, D.** (1997). *WMS-III: Wechsler memory scale administration and scoring manual* (3rd ed.). Psychological Corp.

**West, R.**, & **Alain, C.** (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology*, **37**(2), 179–189. https://doi.org/10.1111/1469-8986.3720179.

**Wheeldon, L. R.**, & **Monsell, S.** (1992). The locus of repetition priming of spoken word production. *The Quarterly Journal of Experimental Psychology*, **44**(4), 723–761. https://doi.org/10.1080/14640749208401307.

**Zec, R. F.**, **Burkett, N. R.**, **Markwell, S. J.**, & **Larsen, D. L.** (2007). A cross-sectional study of the effects of age, education, and gender on the Boston naming test. *The Clinical Neuropsychologist*, **21**(4), 587–616. https://doi.org/10.1080/13854040701220028.

**Zec, R. F.**, **Markwell, S. J.**, **Burkett, N. R.**, & **Larsen, D. L.** (2005). A longitudinal study of confrontation naming in the "normal" elderly. *Journal of the International Neuropsychological Society*, **11**(6), 716–726. https://doi.org/10.1017/S1355617705050897.

https://doi.org/10.1017/S1366728925100229 Published online by Cambridge University Press



## Appendix A

**Table A1.** Lexical properties from the English Lexicon Project database (Balota et al., 2007) for each word in the picture naming task

| Word | Cue | Length | Phonemes | Syllables | Part of speech | Frequency[a] | Orthographic neighbors[b] | Phonological neighbors[b] | Concreteness rating[c] | Semantic neighborhood density[d] | Mean RT in ms[e] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Crown | Worn on your head | 5 | 4 | 1 | *n.* / v. | 8962 | 8 | 11 | 4.81 | 0.65 | 588.75 |
| Helicopter | Used for travel | 10 | 9 | 4 | *n.* | 3201 | 0 | 0 | 4.62 | 0.61 | 593.33 |
| Barrel | A type of container | 6 | 5 | 2 | *n.* / v. | 6836 | 4 | 11 | 4.86 | 0.59 | 625.22 |
| Tiger | A type of animal | 5 | 4 | 2 | *n.* | 5393 | 5 | 8 | 5 | 0.63 | 669.09 |
| Rolling pin | Used for baking | | | | | | | | | | |
| Spool of thread | Used for sewing | | | | | | | | | | |
| Stirrup | Used on a horse | | | | | | | | | | |
| Violin | Musical instrument | 6 | 6 | 2 | *n.* | 2583 | 0 | 1 | 4.96 | 0.59 | 632.11 |
| Iron | Used on clothing | 4 | 3 | 1 | *n.* / v. / adj. | 19187 | 3 | 1 | 4.59 | 0.66 | 565.22 |
| Alligator/ | A type of animal | 9 | 7 | 4 | *n.* | 723 | 0 | 0 | 4.96 | 0.46 | 653.37 |
| Crocodile | | 9 | 8 | 3 | *n.* | 521 | 0 | 0 | 4.83 | 0.53 | 662.11 |
| Pliers | A type of tool | 6 | 5 | 1 | *n.* | 730 | 1 | 5 | 4.93 | 0.3 | 702 |
| Kangaroo | A type of animal | 8 | 7 | 3 | *n.* | 818 | 0 | 0 | 4.86 | 0.5 | 674.77 |
| Gavel | Used by a judge | | | | | | | | | | |
| Duck | A type of animal | 4 | 3 | 1 | *n.* / v. | 6829 | 16 | 37 | 4.86 | 0.6 | 572.86 |
| Guitar | Musical instrument | 6 | 5 | 2 | *n.* | 24781 | 0 | 1 | 4.9 | 0.64 | 631.71 |
| Trombone | Musical instrument | 8 | 7 | 2 | *n.* | 709 | 0 | 0 | 4.9 | 0.57 | 770.17 |
| Well | A source of water | 4 | 3 | 1 | adv. / other/ adj. / *n.* / v. | 552532 | 15 | 40 | 3.33 | 0.7 | 555.04 |
| Rickshaw | A type of transportation | | | | | | | | | | |
| Rhinoceros | A type of animal | 10 | 9 | 4 | *n.* | 117 | 0 | 0 | 4.75 | 0.44 | 793.44 |
| Basket | A type of container | 6 | 6 | 2 | *n.* | 2815 | 4 | 4 | 5 | 0.56 | 584.07 |
| Lobster | A type of animal | 7 | 6 | 2 | *n.* | 849 | 1 | 1 | 4.86 | 0.49 | 638.67 |
| Cummerbund | An article of clothing | 10 | 8 | 3 | *n.* | 28 | 0 | 0 | 4.04 | 0.32 | 812.68 |
| Pipe | Used for smoking | 4 | 3 | 1 | *n.* / v. | 12205 | 7 | 19 | 4.88 | 0.61 | 569.96 |
| Belt | An article of clothing | 4 | 4 | 1 | *n.* | 10820 | 13 | 19 | 4.9 | 0.62 | 575.82 |
| Ostrich | A type of animal | 7 | 6 | 2 | adj. / *n.* | 429 | 0 | 0 | 4.71 | 0.38 | 679.48 |
| Ottoman | A piece of furniture | 7 | 6 | 3 | *n.* | 1576 | 0 | 0 | 4.9 | 0.6 | 724.44 |

(*Continued*)

https://doi.org/10.1017/S1366728925100229 Published online by Cambridge University Press

**Table A1.** (*Continued*)

| Word | Cue | Length | Phonemes | Syllables | Part of speech | Frequency[a] | Orthographic neighbors[b] | Phonological neighbors[b] | Concreteness rating[c] | Semantic neighborhood density[d] | Mean RT in ms[e] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Dresser | Used for storage | 7 | 5 | 2 | *n.* | 874 | 4 | 5 | 4.96 | 0.39 | 634.08 |
| Ruler | Used for measurement | 5 | 4 | 2 | *n.* | 2703 | 4 | 11 | 4.66 | 0.6 | 595.97 |
| Spinning wheel | Used to make thread/ yarn | | | | | | | | | | |
| Asparagus | A type of vegetable | 9 | 9 | 4 | *n.* | 285 | 0 | 0 | 4.96 | 0.4 | 849.22 |
| Candelabra | Used for lighting | 10 | 10 | 4 | *n.* | 2589 | 0 | 0 | 4.88 | 0.25 | 851.67 |
| Leopard | A type of animal | 7 | 5 | 2 | *n.* | 697 | 2 | 8 | 5 | 0.55 | 645.56 |
| Racquet | Used for sports | 7 | 5 | 2 | *n.* | 597 | 0 | 8 | N/A | 0.35 | 687.78 |
| Sheep | A type of animal | 5 | 3 | 1 | *n.* | 7524 | 7 | 32 | 4.9 | 0.6 | 625.21 |
| Doorknob | Found on a door | 8 | 6 | 2 | *n.* | 184 | 0 | 0 | 4.97 | 0.2 | 661.96 |
| Ear | A part of the body | 3 | 2 | 1 | *n.* | 11312 | 12 | 24 | 5 | 0.59 | 630.59 |
| Boot | An article of clothing | 4 | 3 | 1 | *n.* / v. | 43868 | 18 | 43 | 4.96 | 0.58 | 629.81 |
| Ring | A piece of jewelry | 4 | 3 | 1 | *n.* / v. | 44167 | 13 | 31 | 4.81 | 0.66 | 582.58 |
| Grasshopper | A type of insect | 11 | 8 | 3 | *n.* | 400 | 0 | 0 | 4.91 | 0.37 | 658.08 |
| Nail file | Used to smooth | | | | | | | | | | |
| Screwdriver | A type of tool | 11 | 9 | 3 | *n.* | 1173 | 0 | 0 | 4.9 | 0.33 | 819.82 |
| Glasses | Used to enhance vision | 7 | 6 | 2 | *n.* | 8546 | 3 | 3 | 4.9 | 0.57 | 581.96 |
| Record player | Used to play music | | | | | | | | | | |
| Anchor | A heavy device | 6 | 4 | 2 | *n.* / v. | 2766 | 0 | 2 | 4.77 | 0.61 | 627.59 |
| Necklace | Worn around the neck | | | | | | | | | | |
| Pineapple | A type of fruit | 9 | 6 | 3 | *n.* | 637 | 0 | 0 | 4.94 | 0.48 | 651.65 |
| Nut | A fastener | 3 | 3 | 1 | *n.* | 4431 | 17 | 36 | 4.52 | 0.56 | 611.61 |
| Beetle | a type of insect | | | | | | | | | | |
| Bridle | Worn by a horse | 6 | 5 | 2 | *n.* / v. | 380 | 1 | 8 | 4.27 | 0.36 | 730.2 |
| Hanger | Found in closet | 6 | 4 | 2 | *n.* | 961 | 11 | 15 | 4.81 | 0.36 | 623.42 |
| Hammer | A type of tool | 6 | 4 | 2 | *n.* / v. | 6714 | 3 | 9 | 4.77 | 0.61 | 600.41 |
| Barn | A type of building | | | | | | | | | | |
| Abacus | Used for counting | 6 | 6 | 3 | *n.* | 513 | 0 | 0 | 4.52 | 0.3 | 792.69 |
| Eagle | A type of bird | 5 | 3 | 2 | *n.* | 6845 | 1 | 5 | 5 | 0.63 | 656.2 |
| Artichoke | A type of vegetable | 9 | 7 | 3 | *n.* | 181 | 0 | 0 | 4.63 | 0.28 | 677.96 |

(*Continued*)

**Table A1.** (*Continued*)

| Word | Cue | Length | Phonemes | Syllables | Part of speech | Frequency[a] | Orthographic neighbors[b] | Phonological neighbors[b] | Concreteness rating[c] | Semantic neighborhood density[d] | Mean RT in ms[e] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Lightswitch | For turning something on and off | | | | | | | | | | |
| Carafe | Used to hold water | 6 | 5 | 2 | *n.* | 63 | 0 | 3 | 4.6 | 0.19 | 888.07 |
| Eye | A part of the body | 3 | 1 | 1 | *n.* / v. | 41052 | 10 | 17 | 4.9 | 0.66 | 578.54 |
| Mushroom/ | Something to eat | 8 | 6 | 2 | *n.* / v. | 1873 | 0 | 0 | 4.83 | 0.55 | 575.18 |
| Toadstool | | 9 | 7 | 2 | *n.* | 38 | 0 | 0 | 4.82 | 0.19 | 731.63 |
| Wrench | A type of tool | 6 | 4 | 1 | v. / *n.* | 1389 | 4 | 7 | 4.93 | 0.37 | 645.46 |
| Onion | A type of vegetable | 5 | 5 | 2 | *n.* | 2587 | 2 | 0 | 4.86 | 0.51 | 629.07 |
| Centaur | A mythical creature | 7 | 6 | 2 | *n.* | 922 | 0 | 1 | 3.73 | 0.42 | 823.44 |
| Axe | Used for chopping | 3 | 3 | 1 | *n.* / v. | 3070 | 12 | 13 | 5 | 0.56 | 604.75 |
| Nail | Used with a hammer | 4 | 3 | 1 | *n.* / v. | 4603 | 13 | 46 | 4.93 | 0.55 | 593.88 |
| Squirrel | A type of animal | 8 | 7 | 2 | *n.* / v. | 1989 | 0 | 0 | 4.89 | 0.53 | 717.89 |
| Lips | A part of the body | 4 | 4 | 1 | *n.* | 14796 | 14 | 23 | N/A | 0.56 | 588.04 |
| Mitten | An article of clothing | 6 | 4 | 2 | *n.* | 139 | 2 | 8 | 4.89 | 0.2 | 626.04 |
| Cannon | A type of weapon | 6 | 5 | 2 | *n.* | 7579 | 4 | 12 | 4.79 | 0.62 | 634.77 |
| Stroller | Used for pushing a baby | 8 | 6 | 2 | *n.* | 596 | 1 | 2 | 4.96 | 0.23 | 702.59 |
| Gorilla | A type of animal | 7 | 6 | 3 | *n.* | 1346 | 0 | 12 | 4.97 | 0.52 | 629.31 |
| Pomegranate | A type of fruit | 11 | 10 | 4 | *n.* | 91 | 0 | 0 | 4.86 | 0.36 | 890.15 |
| Wagon | Used to pull things | 5 | 5 | 2 | *n.* | 3687 | 0 | 0 | 4.89 | 0.58 | 602.29 |
| Tambourine | A musical instrument | 10 | 8 | 3 | *n.* | 181 | 0 | 0 | 4.86 | 0.42 | 807.56 |
| Heart | A symbol of love | 5 | 4 | 1 | *n.* / v. | 45783 | 4 | 21 | 4.52 | 0.68 | 579.96 |
| Zebra | A type of animal | 5 | 5 | 2 | *n.* | 915 | 0 | 1 | 4.86 | 0.47 | 643 |
| Screw | For holding things in place | 5 | 4 | 1 | *n.* / v. | 9543 | 3 | 5 | 4.81 | 0.56 | 700.14 |
| Celery | A type of vegetable | 6 | 6 | 3 | *n.* | 732 | 0 | 1 | 4.8 | 0.42 | 701.86 |
| Calipers | Used for measuring | 8 | 7 | 3 | *n.* | 322 | 1 | 9 | 4.78 | 0.36 | 829.11 |
| Stool | Used to sit on | 5 | 4 | 1 | *n.* | 1234 | 4 | 17 | 4.9 | 0.43 | 681.3 |
| Seahorse | A type of animal | 8 | 6 | 2 | *n.* | 52 | 0 | 1 | 4.89 | 0.31 | 667 |

(*Continued*)

**Table A1.** (*Continued*)

| Word | Cue | Length | Phonemes | Syllables | Part of speech | Frequency[a] | Orthographic neighbors[b] | Phonological neighbors[b] | Concreteness rating[c] | Semantic neighborhood density[d] | Mean RT in ms[e] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bow | Used to decorate a gift | 3 | 2 | 1 | *n.* | 7503 | 23 | 64 | 4.61 | 0.61 | 631.58 |
| Roller skate | Used in sports | | | | | | | | | | |
| Glove | Worn on a hand | 5 | 4 | 1 | *n.* / v. | 2355 | 3 | 2 | 4.97 | 0.54 | 597.63 |
| Peacock | A type of animal | 7 | 5 | 2 | *n.* | 969 | 0 | 0 | 5 | 0.55 | 676.73 |
| Vest | An article of clothing | 4 | 4 | 1 | *n.* | 1238 | 11 | 22 | 4.52 | 0.46 | 640.66 |
| Kettle | Used to boil water | 6 | 4 | 2 | *n.* | 1499 | 4 | 17 | 4.75 | 0.47 | 655.11 |
| Bunsen burner | It can be lit | | | | | | | | | | |
| Colander | Used for cooking | 8 | 7 | 3 | *n.* | 118 | 0 | 1 | 4.21 | 0.15 | 765.5 |
| Coat | An article of clothing | 4 | 3 | 1 | *n.* / v. | 10295 | 11 | 36 | 4.97 | 0.6 | 562 |
| Trumpet | A musical instrument | 7 | 7 | 2 | *n.* / v. | 7157 | 2 | 1 | 4.86 | 0.57 | 633.19 |
| Trowel | A gardening tool | 6 | 5 | 1 | *n.* / v. | 38 | 0 | 1 | 4.16 | 0.23 | 754.22 |
| Racoon | A type of animal | 7 | 5 | 2 | *n.* | 275 | 0 | 0 | 4.67 | 0.42 | 662.74 |
| Blouse | An article of clothing | | | | | | | | | | |
| Saltshaker | Contains something to season food | 10 | 8 | 3 | N/A | 6 | 0 | 0 | 4.96 | 0.25 | 741.71 |
| Arrow | Provides direction | 5 | 3 | 2 | *n.* | 7324 | 0 | 8 | 4.97 | 0.6 | 573.93 |
| Accordion | A musical instrument | 9 | 8 | 3 | *n.* | 590 | 0 | 0 | 4.86 | 0.56 | 766.15 |
| Pepper | A type of vegetable | 6 | 4 | 2 | *n.* / v. | 5324 | 2 | 8 | 4.59 | 0.57 | 615.14 |
| Broom | Used for cleaning | 5 | 4 | 1 | *n.* | 600 | 4 | 19 | 4.89 | 0.46 | 617.65 |
| Top | A toy | 3 | 3 | 1 | adj. / *n.* / v. / adv. | 116571 | 21 | 27 | 3.93 | 0.68 | 526.43 |
| Pitcher | Used for serving water | 7 | 4 | 2 | *n.* | 1571 | 3 | 6 | 4.93 | 0.57 | 584.52 |
| Chisel | A type of tool | 6 | 4 | 2 | *n.* | 2697 | 0 | 3 | 4.63 | 0.35 | 750.85 |
| Metronome | Used in music to keep time | 9 | 8 | 3 | *n.* | 411 | 0 | 0 | 4.27 | 0.3 | 875.79 |
| Sled/ | Used to play in the snow | 4 | 4 | 1 | *n.* | 2423 | 9 | 13 | 5 | 0.45 | 669.37 |
| Sleigh | | 6 | 3 | 1 | *n.* | 739 | 0 | 19 | 4.71 | 0.35 | 689 |
| Hand | A part of the body | 4 | 4 | 1 | *n.* / v. | 124848 | 11 | 17 | 4.72 | 0.69 | 578.19 |
| Monocle | Used to magnify | 7 | 6 | 3 | *n.* | 354 | 1 | 1 | 3.87 | 0.26 | 704.13 |

(*Continued*)

https://doi.org/10.1017/S1366728925100229 Published online by Cambridge University Press



**Table A1.** (*Continued*)

| Word | Cue | Length | Phonemes | Syllables | Part of speech | Frequency[a] | Orthographic neighbors[b] | Phonological neighbors[b] | Concreteness rating[c] | Semantic neighborhood density[d] | Mean RT in ms[e] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Thimble | Used when sewing | 7 | 5 | 2 | *n.* | 114 | 1 | 4 | 5 | 0.24 | 761.88 |
| Corn | A type of vegetable | 4 | 4 | 1 | *n.* | 4988 | 17 | 32 | 4.96 | 0.59 | 621.11 |
| Clothspin | Used to hang things | 10 | 7 | 2 | *n.* | 151 | 0 | 0 | 4.41 | 0.19 | 866.33 |
| Chicken | A type of animal | 7 | 5 | 2 | *n.* / v. | 11478 | 1 | 4 | 4.8 | 0.59 | 686.43 |
| Harp | A musical instrument | 4 | 4 | 1 | *n.* / v. | 1667 | 9 | 11 | 4.85 | 0.56 | 575.21 |
| Pumpkin | A type of vegetable | 7 | 7 | 2 | *n.* | 937 | 1 | 1 | 4.9 | 0.47 | 573.31 |
| Watering can | Used in the garden | | | | | | | | | | |
| Saw | Used by a carpenter | 3 | 2 | 1 | v. / *n.* | 87044 | 20 | 32 | 4.46 | 0.69 | 609.85 |
| Dragonfly | A type of insect | 9 | 9 | 3 | *n.* | 1856 | 0 | 0 | 4.83 | 0.39 | 706.29 |
| Pear | A type of fruit | 4 | 3 | 1 | *n.* | 971 | 17 | 45 | 4.93 | 0.47 | 658.64 |
| Rocking chair | A piece of furniture | | | | | | | | | | |
| Windmill | Used to convert energy | 8 | 7 | 2 | *n.* | 302 | 0 | 0 | 4.89 | 0.54 | 607.77 |
| Butterfly | A type of insect | 9 | 7 | 3 | *n.* / adv. / adj. | 1798 | 0 | 0 | 4.93 | 0.57 | 620.26 |
| Flute | A musical instrument | | | | | | | | | | |

*Note*: The following items were not included in the analyses of this study: rolling pin, spool of thread, stirrup, gavel, rickshaw, spinning wheel, nail file, record player, necklace, beetle, barn, ironing board, Bunsen burner, blouse, watering can, rocking chair and flute. The items alligator/crocodile, mushroom/toadstool and sled/sleigh had two correct responses and therefore have descriptive characteristics for both options. *n.* = noun, v. = verb, adj. = adjective, adv. = adverb, RT = Reaction Time.
[a]Frequency norms are from the Hyperspace Analogue to Language (HAL) corpus (Lund & Burgess, 1996).
[b]Orthographic and phonological neighborhood sizes represent the number of words differing by one letter or one phoneme, respectively, excluding homophones (Balota et al., 2007).
[c]Concreteness ratings range from 1 (abstract) to 5 (concrete) (Brysbaert et al., 2014).
[d]Semantic neighborhood density is calculated using the number of words that are close in meaning surrounding a target word (see Shaoul & Westbury, 2010 for a detailed explanation of calculations).
[e]Mean reaction time (in milliseconds) naming data are from the English Lexicon Project (Balota et al., 2007).