Applied Psycholinguistics (2026), **47**, e59, 1–38
doi:10.1017/S0142716426100903

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

SpaVerb-LD: A megastudy of Spanish verb processing and lexical-semantic dynamics

María González-Nosti[1], Javier Marín[2], Romina San Miguel-Abella[1], Elena Herrera[1] and Miguel Angel Pérez-Sánchez[2]

[1]Psicología, Universidad de Oviedo, Spain and [2]Psicología Básica y Metodología, Universidad de Murcia, Spain
Corresponding author: María González-Nosti; Email: gonzaleznmaria@uniovi.es

(Received 29 July 2025 UTC; revised 7 May 2026 UTC; accepted 14 July 2026 UTC)

Abstract
Understanding how verbs are processed during lexical access remains an open theoretical question, as most large-scale evidence has focused on nouns. This megastudy analyzes variables involved in the processing of Spanish verbs through response times obtained in a lexical decision task. The dataset, named SpaVerb-LD (Spanish Verb Lexical Decision), continues the SpaVerb-WN (Spanish Verb Word Naming), allowing direct comparison across tasks. The linear mixed-effects model revealed consistent facilitatory effects of word frequency and neighborhood size and inhibitory effects of word length and age of acquisition (AoA). For the first time in Spanish, we observed a facilitatory effect of word prevalence and an inhibitory effect of verb motor content. Previously reported two-way interactions were also replicated, including those between word prevalence and frequency, prevalence and neighborhood size, and AoA and length. A paradoxical interaction between AoA and word frequency was also identified. Crucially, these interactions indicate that lexical decision performance is shaped by the interplay between lexical and semantic variables, highlighting the role of semantic richness in verb processing. The findings support the view that semantic access is involved in lexical decision even in transparent languages and demonstrate that word prevalence and motor content are key to understanding lexical-semantic verb processing beyond traditional frequency-based accounts.

Keywords: lexical decision; megastudy; motor content; verbs; word prevalence

Introduction

For decades, psycholinguistic research has focused, among other topics, on the impact of stimuli's features on reading, as they allow inferences to be made about the cognitive processes involved in reading skills and their characteristics. Importantly, differential effects have been observed between languages, particularly those that differ in the orthographic transparency and metric systems (Ardila, 1991; Ardila &

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Cuetos, 2016). Most of the evidence regarding word recognition and reading comes from empirical studies conducted in English. Consequently, numerous influential models of word recognition and reading are built upon findings from English (for a comprehensive overview, refer to Norris, 2013). However, several linguistic characteristics unique to English, not found in other languages, could significantly impact word processing. For instance, English features an opaque orthographic system characterized by highly inconsistent grapheme–phoneme correspondences, lacks consistent accentuation rules, and has ambiguous syllabic segmentation. In contrast, other languages possess transparent orthographic systems with greater consistency in grapheme–phoneme mappings, clear accentuation rules, and distinct syllabic boundaries. Spanish, for example, exemplifies such characteristics, making it an appropriate language for comparison with English. Notably, as elaborated further, evidence indicates that the effects of various psycholinguistic variables associated with lexical items (i.e., words) differ across languages. Hence, conclusions concerning the processing of printed words in one language cannot be directly extrapolated to others. Thus, adjustments would be required for English-based reading models when applied to other languages, accounting for the presence and influence levels of the aforementioned psycholinguistic variables.

From a theoretical perspective, the present study is framed within the models of lexical access that differ in the role attributed to semantic information during word recognition. Interactive activation models (e.g., Coltheart et al., 2001) assume bidirectional interactions between orthographic, phonological, and semantic levels, allowing semantic information to influence early stages of processing. In contrast, frequency-based and distributional approaches emphasize the role of exposure-driven lexical representations, often minimizing the necessity of semantic access for tasks such as lexical decision (LD) (e.g., Norris, 2006). Additionally, embodied or grounded cognition frameworks propose that word meaning—particularly for action-related words such as verbs—is partly represented in sensorimotor systems (e.g., Pulvermüller, 2005, 2013).

Within this theoretical landscape, a key unresolved question concerns whether semantic access is necessarily engaged during lexical decision, especially in transparent orthographies such as Spanish, where sublexical processing may suffice for accurate performance. The present study directly addresses this issue by examining the extent to which lexical-semantic variables, including word prevalence and motor content, contribute to verb recognition performance.

Methodologically, studies investigating variables that impact processing have evolved from factorial designs with a limited number of stimuli and very specific characteristics, to megastudies, which offer advantages over traditional factorial design experiments in cognitive psychology in general and particularly in psycholinguistics, as noted in numerous previous works (see e.g., Balota et al., 2012; Keuleers & Balota, 2015; Kuperman, 2015). Basically, megastudies allow us to explore or confirm the contribution of a large number of factors to behavioral data and ensure sensibly the generalizability of findings to an entire language. Despite these advantages, megastudies also present some limitations, such as reduced experimental control over specific variables, potential heterogeneity in participant samples (especially in large-scale online data collection), and challenges in

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

establishing causal relationships due to their largely correlational nature (Keuleers & Balota, 2015; Kuperman, 2015).

Since the pioneering megastudies by Coltheart (1981) and by Seidenberg and Waters (1989), many others have been conducted in different languages. The iconic English Lexicon Project (Balota et al., 2007), with chronometric data of word naming and visual lexical decision (i.e., a task in which participants decide whether letter strings presented on a screen are real words or pseudowords) for 40,481 American English words, had their equivalents in the lexical decision task in Dutch (Keuleers et al., 2010; with 14,000 words), French (Ferrand et al., 2010; with 38,840 words), British English (Keuleers et al., 2012; with 28,730 words), and Spanish (Aguasvivas et al., 2018; SPALEX, with over 45,000 words). These megastudies, along with others, are listed on the Center for Reading Research's website (Brysbaert, 2022). SPALEX (Aguasvivas et al., 2018), a relevant reference for our current megastudy, collected data from an online lexical decision task, which was performed by nearly 170,000 Spanish native speakers (who reported Spanish as their first-acquired and dominant language) from Spain and Spanish-speaking countries of America. Participants had to respond to 100 items (70 words and 30 pseudowords) with no time limit imposed on individual responses. This database has been used to examine the influence on vocabulary size (Aguasvivas et al., 2020), and to investigate age-related differences in the recognition of valenced words (Kyröläinen et al., 2021). Nevertheless, response time (RT) data from SPALEX have not been thoroughly analyzed to provide a comprehensive examination of the psycholinguistic variables implicated in word recognition in Spanish. Other notable megastudies have also been conducted in Spanish, including word naming (Davies et al., 2013), lexical decision (González-Nosti et al., 2014), and more recent work focusing on verbs (SpaVerb-WN; San Miguel-Abella et al., 2022).

Against this background, the present work extends this line of research by conducting a megastudy on visual lexical decision in Spanish for the same set of verbs. Word naming (or reading aloud) and visual lexical decision are two commonly used cognitive tasks in reading research. Although some variables consistently show similar effects in both tasks (in direction and size) across languages (e.g., lexical frequency, with lower RT and higher accuracy for high-frequency than for low-frequency words), others yield mixed results depending on the tasks (e.g., neighborhood size) and languages (e.g., syllabic frequency). These similarities and differences are typically interpreted in terms of shared and task-specific processes involved, as well as the orthographic transparency of the language. In addition, some variables, such as the motor content of words, have received comparatively little attention in lexical decision tasks and warrant further investigation.

Psycholinguistic effects in lexical decision

In visual lexical decision tasks, word frequency has consistently shown a significant influence on word recognition. High-frequency words tend to have better accuracy rates and require less time to be recognized than low-frequency words, regardless of the language being studied (for a review, see Ghyselinck et al., 2004), including those with a highly transparent orthography such as Spanish (e.g., González-Nosti et al.,

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

2014) or Dutch (e.g., Brysbaert et al., 2016) and even those with a full transparent orthography, as is the case of Persian (Nemati et al., 2022).

Word length (number of letters) is another variable that directly affects word recognition. In English, a quadratic effect of word length on word recognition times has been shown when word frequency is controlled (i.e., RTs decrease for two-to-four letter words, stay stable for five-to-eight letter words, and increase for longer words; see New et al., 2006). A similar U-shaped effect of word length has been observed in languages with intermediate-depth orthography as French (e.g., Ferrand et al., 2010) or Portuguese (Soares et al., 2019). However, in more transparent languages the effect seems to be linear, as in Dutch (Brysbaert et al., 2016), or is not significant, as in Persian (Nemati et al., 2022), Italian (Colombo & Burani, 2002), and Spanish (González-Nosti et al., 2014). This null effect of word length on recognition times in transparent languages aligns with the findings of Acha and Perea (2008), in Spanish, where the inhibitory effect of word length found in children (i.e., long words were read slower than short words) disappears in adults. These authors thus reported a developmental shift from an inhibitory length effect in RTs to a null effect in skilled adult readers. They further suggested that in a literate adult lexical system, at least in a transparent language, access to known word representation is automatic, whereas accessing unknown words or pseudowords requires a letter-by-letter decoding (see also Barton et al., 2014).

In contrast, while considering accuracy rather than RTs, Aguasvivas et al. (2020) demonstrated that word length was a significant facilitatory predictor of accuracy rate in lexical decision in Spanish, such that longer words were recognized more accurately than shorter ones. Unfortunately, Aguasvivas et al. did not analyze the influence of word length (or any other variable) on RTs, and González-Nosti et al. did not analyze the length effect (nor any other variable) on accuracy.

This contrast across studies highlights a discrepancy between dependent measures. High accuracy rates and low RT both indicate ease of word recognition. However, the direction of the length effect differs depending on the measure; the facilitatory main effect of word length on accuracy, observed by Aguasvivas et al. in Spanish and previously by Yap and Balota (2009) in English, means that the longer the words are, the easier (more accurate) they are recognized, whereas the robust inhibitory effect on RT found in the languages mentioned above goes in the opposite direction because the greater the length, the more difficult (longer RT) is the correct recognition. At present, there is no satisfactory explanation that reconciles these opposing patterns across measures.

With respect to interactions, Aguasvivas et al. (2020) also identified an interaction effect between word length and frequency in accuracy; a clear and direct impact of length was observed for low-frequency words (i.e., longer words exhibited higher accuracy), an effect that gradually diminished as frequency increases. Similar interactions between these two variables have been reported in many previous studies on word recognition times in English and other languages (see Barton et al., 2014, for review). However, in contrast to these prior findings, this interaction was not found in the megastudy by Gonzalez-Nosti et al. and, as far as we know, it has not been previously reported in Spanish.

The orthographic neighborhood of words has been observed to result in a facilitative impact on word recognition. Words with larger neighborhoods tend to

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

elicit more accurate responses and quicker recognition times. This phenomenon is commonly explained by the hypothesis that words with dense neighborhoods activate numerous lexical representations. As a result, the "word-response" can rely on the total activation in the lexicon rather than solely on the activation of the individual word representation. This effect has been consistently observed across languages, with robust effects for opaque languages as English (e.g., Andrews, 1997; Yap & Balota, 2009), for intermediate-depth orthography languages as French (e.g., Ferrand et al., 2010) or Portuguese (Soares et al., 2019), and for transparent languages as Spanish (e.g., Aguasvivas et al., 2020; González-Nosti et al., 2014), but see the mixed results found by Nemati et al. (2022) in Persian. Orthographic neighborhood effect seems to interact with the effect of other variables. The interaction between neighborhood and word frequency, so that the facilitatory effect of the former is higher for low-frequency words than for high-frequency words, has been widely reported across languages (Andrews, 1997; Brysbaert et al., 2016; Yap & Balota, 2009). In Spanish, Gonzalez-Nosti et al. did not find this interaction as significant while predicting RTs, and Aguasvivas et al. (2020) found the interaction as significant effect on accuracy, but in the opposite way, with a notable inhibitory effect of neighborhood in low-frequency words and a very small effect in high-frequency words. The interaction between orthographic neighborhood and word length has also been reported. The neighborhood effect seems to be only significant for short words (e.g., Brysbaert et al., 2016; González-Nosti et al., 2014). This association between length and neighborhood is predictable, as shorter words tend to have denser neighborhoods, whereas longer words typically have sparser neighborhoods. This interaction may be considered a trivial effect, as it arises from the inherent relationship between word length and neighborhood size. Alternatively, it has been suggested that the advantage of having a dense neighborhood (i.e., of having many similar words that contribute to an overall activation in the lexicon) could take out the effects of frequency or length (Brysbaert et al., 2016).

Age of acquisition (AoA) is a decisive variable in visual lexical decision, with early-acquired words recognized faster than late-acquired ones (see Elsherif et al., 2023; Johnston & Barry, 2006; Juhasz, 2005 for reviews). This effect has been observed across different languages when analyzed, mostly in English (e.g., Brysbaert & Cortese, 2011; Cortese & Khanna, 2007; Cortese & Schock, 2013; Morrison & Ellis, 2000), but also in French (Bonin et al., 2001), Portuguese (Soares et al., 2019) Dutch (Brysbaert et al., 2000; Ghyselinck et al., 2004), Italian (Colombo & Burani, 2002; Menenti & Burani, 2007), and Spanish (González-Nosti et al., 2014; Wilson et al., 2013). Although it has been an intense debate about how genuine the AoA effect is, as it could be confounded with the effects of other variables such as cumulative frequency and frequency of trajectory, current evidence strongly suggests certain and separate effects of AoA in word recognition (Izura et al., 2011; Juhasz et al., 2019; see also Brysbaert, 2017). Additionally, an interaction between AoA and word frequency has been observed in both opaque (Cortese & Schock, 2013) and transparent languages (Brysbaert et al., 2016; González-Nosti et al., 2014). This interaction indicates that the effect of AoA is stronger or exclusively present in low-frequency words. Additionally, there is an interaction between AoA and word

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

length where the AoA effect tends to be stronger for long words (Brysbaert et al., 2016; González-Nosti et al., 2014).

Regarding syllable frequency, the seminal work by Carreiras et al. (1993) was the first to demonstrate an effect of the first-syllable frequency in a visual lexical decision in Spanish because words starting with high frequency syllables were recognized slower than those starting with low frequency syllables (see also Carreiras & Perea, 2004). This inhibitory effect was ascribed to the influence of syllable frequency on the recognition process. Specifically, words beginning with high-frequency syllables activate a broader range of potential candidates compared with those starting with low-frequency syllables. Then, the larger the number of activated candidates, lexical selection processes take longer. The effect has been replicated in French (e.g., Chetail & Mathey, 2009b) and German (e.g., Conrad et al., 2006), but a later study in Spanish failed in observing behavioral effects of syllable frequency on accuracy and RT in a visual lexical decision task (Carreiras et al., 2006). In other lexical decision experiments, a dissociation between type and token measures of syllable-frequency was found, with the token frequency predicting an inhibitory effect on word recognition times, and with the type frequency predicting a facilitation effect (Conrad et al., 2008). Interestingly, the inhibitory effects mentioned above have not been replicated in English, where the effect of syllable frequency in lexical decision showed as facilitatory instead (i.e., words starting with high-frequency syllables are recognized faster compared to those beginning with low-frequency syllables; Macizo & Van Petten, 2007).

Word prevalence, defined as the percentage of individuals who know a word, is a relatively new variable (it was presented by Keuleers et al., 2015, and it is calculated from accuracy in a lexical decision task, with corrections for response bias due to random word recognition) that has demonstrated to play a prominent role in word processing, surpassing even the effects of word frequency, AoA, and length. Despite those remarkable results, prevalence has been only analyzed in word naming and visual lexical decision tasks with Dutch (Brysbaert et al., 2016) and English words (Brysbaert et al., 2019; see also for an extensive analysis of the different roles of word prevalence). Brysbaert et al. (2016, 2019) suggested that word prevalence can be considered a variable that works akin to what a subjective frequency does, a measure to compensate for poor estimates of word frequency based on text corpus counts. Therefore, word prevalence effect could be interpreted as the same way as the frequency effect. Importantly, word prevalence interacts with numerous variables, namely, word frequency, AoA, length, and neighborhood size. In a series of two-way interaction analyses, Brysbaert et al. (2016) showed that the prevalence effect is stronger for mid-frequency, short, late acquired, and "hermit" (i.e., those with sparse orthographic neighborhoods) words. Beyond these results from word recognition in Dutch, these relationships have not been addressed further.

Finally, the role of grammatical category in visual lexical decision has received less attention compared with other variables, but studies have found some influence of grammatical class on performance. Recently, Brysbaert et al. (2016) observed that closed-class words (functional words and numbers) took longer recognition times than open-class words (nouns, verbs, and adjectives) in Dutch. Within the open-class words, faster recognition latencies for nouns than for verbs have been found (e.g., Brysbaert et al., 2012; Sereno & Jongman, 1997, in English; Cordier et al., 2013,

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

in French). Regarding this noun advantage, the experiments in German by Kauschke and Stenneken (2008) demonstrated that the processing differences between nouns and verbs cannot be reduced to variables such as frequency, word form, or morphological complexity. Chiarello et al. (2002), using English words, showed that the noun advantage disappears when nouns and verbs were explicitly matched for imageability. Therefore, it has been suggested that the noun advantage is more related to semantic dimensions than lexical or morphological aspects. However, Colombo and Burani (2002) reported larger effect sizes of AoA (a lexical-semantic variable), and root frequency (a lexical variable), on recognition times for Italian verbs compared to nouns. Although megastudies include words from various grammatical categories, the analysis of the influence of grammatical category in word recognition is still limited (but see Brysbaert et al., 2016). In Spanish, Davies et al. (2013) only provided evidence that verbs and nouns were pronounced significantly slower than adjectives in word naming; González-Nosti et al. (2014) did not include the grammatical category factor in the analysis of lexical decision times and, up to date, there are no studies based on the Spalex Project (Aguasvivas et al., 2018) about this factor. Considering the limited and varied evidence currently available, the question arises naturally as to whether the findings obtained with nouns (such as the facilitatory effect of imageability, the noun advantage in recognition latencies, and the influence of semantic dimensions on processing) can be directly applied to words of other grammatical categories, particularly verbs. Nonetheless, some insights can be gleaned from neuropsychological research, which may supplement the mentioned behavioral evidence with neuropsychological dissociations in verb/noun processing (see the next section).

As observed, the impacts of the aforementioned variables on visual word recognition can vary between languages. Specifically, the effects observed in a language with opaque orthography, like English, should not be automatically applied to languages with transparent orthographies, such as Spanish, and vice versa. Thus, previous study findings indicate the necessity for further research with a concentrated emphasis on the orthographic systems of diverse languages. In line with this, we introduce a new megastudy utilizing extensive data from a visual lexical decision task conducted in Spanish, examining the effects of various psycholinguistic variables on the given task.

Neural evidence on action words processing

Although the primary focus of this study is behavioral, evidence from neural studies of action words can help contextualize how motor content influences verb recognition. There is a consensus among researchers regarding the role of the neural circuits of sensory and motor areas in the organization of the semantic system and in the processing of word meaning. Thus, evidence indicates that categorical dissociations between verbs and nouns are likely rooted in semantic distinctions, with verbs more associated with sensory/action-motor features and nouns with visual/abstract features (Arévalo et al., 2007; Gainotti, 2004; Kiefer & Pulvermüller, 2012). More precisely, there exists a connection between the processing of action words and the motor regions of the brain, which are responsible for planning and executing the corresponding actions (for a review, see Pulvermüller, 2013).

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

This theoretical framework, called "embodiment" or "semantic grounding," states that the semantics of action words (i.e., essentially verbs) are correspondingly embodied with the sensorimotor neural circuits activated during the performance of those actions (Glenberg, 2015; Pulvermüller, 2005). Furthermore, this theory also claims that even abstract concepts are based on the experience (i.e., actions, emotions, and perceptions) of concrete situations, and therefore they have a brain trace at the sensory, motor, and emotional structures. Embodiment theory has increasingly been supported by evidence in both neuroimaging and behavioral studies using language comprehension and production tasks. We will focus on some of the most relevant results found in visual lexical decision. Pulvermüller et al. (1999) examined whether different cortical areas are activated during the recognition of verbs and nouns. They observed the effects of word class on behavioral measures, on event-related potentials, and on high-frequency electro cortical responses. Physiological activity maps derived from both event-related potentials and high-frequency records revealed strong differences between word class over motor and visual cortices. Although there was not an effect of word class on recognition accuracy or latency, participants rated verbs as more motor-related words (i.e., those referred to activities that participants could perform themselves) than nouns, and nouns as more visual-related (i.e., those referred to visually perceivable objects or scenes). Pulvermüller et al. interpreted these physiological responses as reflecting semantic associations, with verbs eliciting motor associations and nouns eliciting visual associations. Moreover, several neuroimaging studies, focusing on nouns related to actions performed with specific body parts, have further supported the involvement of the sensorimotor system in conceptual and semantic language processing (e.g., Carota et al., 2012) As a matter of fact, words referring to actions activate motor areas corresponding to the part of the body involved in performing that action, suggesting distributed cortical processing of semantics (see Kiefer & Pulvermüller, 2012, for a review). Convergent results have been found with transcranial magnetic stimulation on the motor cortex (Pulvermüller et al., 2005) and from behavioral studies. Sidhu et al. (2014) collected ratings of imageability and "relative embodiment," defined as the degree to which the meaning of each verb involves the human body, for a set of verbs. Then, 400 of them jointly with the same number of pseudowords were selected as materials for a lexical decision task. These variables were then examined as predictors of verb recognition latencies in a lexical decision task, while controlling for other psycholinguistic factors. A hierarchical regression analysis revealed that, after the introduction of other psycholinguistic variables as controls, only relative embodiment, and not imageability, emerged as a significant predictor of verb recognition latencies, with faster times for high-embodied than for low-embodied verbs. Similar findings were reported in the study by Sidhu et al. (2016), in which AoA, valence, arousal, semantic diversity, imageability, and relative embodiment were examined as semantic predictors of verb recognition latencies in a lexical decision task. They found that, after applying controls for word frequency and neighborhood, relative embodiment, AoA, valence, and arousal (but not imageability and semantic diversity) were the significant predictors. Taken together, all this evidence indicates a strong linkage between the neural foundation of verb/action representation and the anterior cortical motor systems.

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

The results obtained with healthy individuals have also been confirmed by dissociations from patients with motor cortex involvement, either due to lesion, stroke, or neurodegenerative disease. Neininger and Pulvermüller (2003) asked for patients affected in the right frontal lobe or in the right temporo-occipital lobes to perform an accelerated lexical decision task with different word classes: action verbs, nouns with strong visual associations, and nouns with both action and visual associations. They found that patients with frontal lesions showed more problems in the recognition of action verbs compared to the other conditions, whereas patients with temporal lesions performed worse on visually related nouns. Similar dissociations were found in patients with motor neuron disease, where selective impairment of verb comprehension and production, compared with those of nouns, was observed (Bak et al., 2001). However, former discussion does not apply to all verbs (and to all nouns either). Considering the semantics of the verbs, those with concrete or abstract meanings are processed in partially differentiated brain networks. Patients with lesions in frontal areas (i.e., those related to sensory and motor processing) had problems when named action verbs but not with abstract verbs, although patients with lesions in temporal areas (i.e., those related to semantic processing) had problems naming abstract but not action verbs (Riccardi et al., 2019; Rüschemeyer et al., 2007). Patients with Parkinson's disease also exhibit difficulties in the processing of action verbs that require more movement (i.e., high motor content) compared with those that require less movement (i.e., low motor content) (Bocanegra et al., 2017; Fernandino et al., 2013; Herrera et al., 2012). Studies with patients with aphasia and impaired access to knowledge about manipulable objects, such as tools, showed a selective deficit for both nouns and verbs related to concepts involving action (Arévalo et al., 2007).

The development of new databases and resources that capture the motor and sensorimotor content of words has indeed been valuable in conducting more precise experiments and further investigating the theoretical framework of embodiment. Thus, San Miguel-Abella and González-Nosti (2020) elaborated a database with 4565 Spanish verbs that includes information about their motor content. Lynott et al. (2020) have also developed a sensorimotor database for 40,000 English words, from which information is available for six perceptual modalities (touch, hearing, smell, taste, vision, and interoception) and five action effectors (mouth/throat, hand/arm, foot/leg, head excluding mouth/throat, and torso) including words from all major syntactic categories (e.g., nouns, verbs, adjectives, adverbs, prepositions, pronouns, etc.). In addition, they examined the sensorimotor variables together with data from the English Lexicon Project (Balota et al., 2007) and found that these variables are predictors of both RT and accuracy in lexical decision tasks. In Spanish, the SpaVerb-WN database (San Miguel-Abella et al., 2022) provides RT data for 4,562 Spanish verbs in a word-naming task. The results revealed that RT increases with increasing word length and with decreasing neighborhood size. Additionally, AoA was found to mediate the effect of motor content (i.e., with an inhibitory effect of motor content for early-acquired words and a facilitatory effect for late-acquired words) and word frequency (i.e., the AoA effect diminishes gradually for low-frequency verbs as frequency increases.). Notably, no significant effect of first syllable frequency was observed.

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Beyond theoretical inquiries into verb and noun processing, the present study provides an extensive compilation of word recognition latencies, which holds inherent interest. It encompasses norms for lexical-semantic, lexical, and sublexical variables, alongside new data on word prevalence, motor content, accuracy, and chronometric records for a substantial array of verbs and pseudoverbs. These datasets could streamline stimulus selection processes, facilitate the design of new virtual or pilot experiments, or enable exploration of novel research questions for fellow researchers. Moreover, the open accessibility of this data serves as a valuable resource for scientists across fields such as language processing, attention, memory, and even for the development of computer models of word recognition.

At the clinical level, this database could also prove invaluable, particularly in the selection of materials for rehabilitation. RTs can pinpoint words that are recognized more easily, thereby aiding in the gradation of material difficulty for the treatment of language impairments.

Thus, within this context and as a natural extension of SpaVerb-WN, the primary aim of the present megastudy, SpaVerb-LD, is to investigate whether verb processing engages lexical-semantic mechanisms beyond those predicted by frequency-based or sublexical accounts. Additionally, we sought to delve deeper into the role of verb motor content in word recognition, as well as the interactions between this variable and each of the selected psycholinguistic variables. Additionally, we intend to test some significant effects previously reported in the literature. To achieve this purpose, we used data from a lexical decision megabase (Pérez-Sánchez et al., 2026) and performed several analyses of logistic and linear mixed-effects regression models to address whether the "classical" variables of word frequency, AoA, length, neighborhood size, and syllable frequency, as well as, for the first time in Spanish, word prevalence, motor content, and a series of interactions between the mentioned variables, affect verb recognition accuracy and latencies.

Method

Participants

A total of 267 participants (aged $M = 21.1$, SD $= 3.8$; and 84% were women) performed a lexical decision task. All of them were graduates or undergraduates at the Faculty of Psychology of the University of Murcia or of the University of Oviedo (Spain) and participated in exchange for academic credit or as volunteers. Although participants had received formal instruction in a second language (typically English) from an early age and some had the option to learn a third language at the beginning of secondary education, all were considered native speakers of Spanish based on their current dominance in the language, daily use in communication and academic activities, and literacy in Spanish. Participants not meeting these criteria were excluded. The variable level of proficiency in the second language did not affect their classification as native speakers. A more detailed description of the sample and a simulation-based power analysis can be found in the study of Pérez-Sánchez et al. (2026).

Results from that power analysis confirmed as valid the commons practices in the field of taking over 20–25 responses per item, as a minimum, in megabases with

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

thousands of items. The megabase Spaverb-LD (Pérez-Sánchez et al., 2026) collects a total of valid 248,871 observations. In accordance with the guideline provided by Brysbaert and Stevens (2018), which suggests a minimum of 1,600 observations per condition in designs involving repeated measures, and estimating a loss of 20% of them after removing errors and RT outliers (resulting in around 199,000 responses), we could construct a prediction model for RT of up to 64 parameters with 1,600 observations per parameter.

Materials

We used the set of verbs included in the SpaVerb-LD megabase by Pérez-Sánchez et al. (2026). This compilation comprises 4,565 Spanish verbs in either infinitive or pronominal forms (i.e., infinitive + the reflexive pronoun "-se"). The entire verb set also had prior norms for AoA (Alonso et al., 2016) and motor content (San Miguel-Abella & González-Nosti, 2020). AoA values correspond to the mean of participants' ratings of the age at which each verb was learned, on a scale from 1 (<2 years), 2–10 (aged 2–10 years), and 11 (≥11 years). Motor content values correspond to the mean of participants' ratings of the degree of bodily movement implied by each verb, on a 1–7 scale (1 = low motor content; 7 = high motor content). We gathered data on word length (in number of letters and syllables), frequency (word frequency per million), and neighborhood (through three measures: the number of phonological and orthographic substitution neighbors and the orthographic Levenshtein's distance) from EsPal database (Duchon et al., 2013), a comprehensive, web-accessible repository of Spanish word properties, built from large written and subtitle corpora. First-syllable frequency measures were obtained from San Miguel-Abella et al. (2022), based on the syllabified Real Academia Española CREA corpus (~170 million words). Syllabification followed a rule-based algorithm for Spanish (Cuayáhuitl, 2004). Frequencies were computed for individual syllables (not structural types), with type frequency defined as the number of distinct words beginning with a given syllable and token frequency as the summed frequency of those words. Word prevalence data were obtained from Pérez-Sánchez et al. (2026) and computed following Keuleers et al. (2015) using item response theory to correct for response bias (false positives) in a yes/no word recognition task.

Pseudoverbs were constructed using a semiautomatic multistep procedure based on trigram frequencies. Type and token trigram frequencies were computed from the set of 4,565 verbs used in the study. For each source verb, an initial pool of pseudoverb candidates was generated by concatenating mutually compatible trigrams derived from that corpus. Next, for each candidate, we computed its Euclidean distance from the profile of the paired source verb. This distance measured was used to rank candidates by their distributional proximity to the source item and thus to identify a restricted set of plausible verb-like candidates. Importantly, this distance metric was not used as an automatic final decision rule. Rather, the final pseudoverb paired with each verb was selected by the researchers from the candidate set on the basis of its overall suitability for the task, prioritizing items judged to be well formed and sufficiently verb-like. Accordingly, the present study relied on a semiautomatic stimulus-construction procedure rather than on a fully autonomous pseudoword generator. A more extended technical description of

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

the candidate-generation procedure is provided by Pérez-Sánchez et al. (2026).The selected verbs and pseudoverbs were randomly divided into 18 blocks with 506 or 508 items each (50% of verbs and 50% of pseudoverbs), which were randomly distributed among the different participants.

Procedure

A detailed description of the procedure is reported in the work of Pérez-Sánchez et al. (2026). Essentially, they run a lexical decision task in the standard lab conditions. The task was performed at the Faculties of Psychology at the University of Oviedo or University of Murcia. The presentation of items and recording of responses were conducted using the DMDX software (Forster & Forster, 2003). Data collection occurred during individual sessions, where each participant was sequentially presented with the verbs in lowercase letters, using Arial font at 11-point size, with black characters on a white background. Participants were instructed to respond with speed and accuracy and always pressed the "word" key with their dominant hand and the "pseudoword" key with the other. Each experimental trial started with an asterisk as a fixation point at the center of the screen for 500 ms, and after that an item appeared and remained on the screen until the participant pressed a key or until 1,500 ms. Each participant was presented with the target items in a randomized order, with breaks every 50 stimuli. Each subject performed a maximum of one block of items per day. Each block was presented to at least 25 participants.

Results

We first report the data selection and description for analysis. We then report some basic statistics for response accuracy and speed, and for the characteristics of the psycholinguistic variables. We then present the pairwise correlations matrix of variables and the steps taken to avoid the potential problem of multicollinearity. Finally, we report the results from several linear mixed-effects model (LMM) analyses carried out on word recognition accuracy and latencies.

Data selection and description

We used the behavioral data from the lexical decision megabase by Pérez-Sánchez et al. (2026), which is available at https://osf.io/5cpsh/. For accuracy analyses, we used the complete set of records for the verbs (column "ACC" in the file "SpaVerb-LD_raw_database"). This set is composed of a total of 124,134 responses from 267 participants to the 4,565 verbs, grouped in 491 individual records, and with a minimum of 24 different responses per item. Each response was coded as "1" (correct) or "0" (incorrect). For response RT analyses, we used the trimmed dataset (column "RT_filtered" in the file "SpaVerb-LD_RT_filtered"), which is composed of only correct responses and where RT outliers are removed. Outliers were detected and removed by applying commonly cut thresholds used in psycholinguistics. First, on the raw data records, the mean RTs (in ms) of the correct responses were calculated by each individual record. Then, the responses with extreme fast

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Table 1. Descriptive statistics of the accuracy rate, RT, and the predictor variables

	N	Mean	SD	Median	Min	Max
Accuracy rate (verbs)	4565	0.81	0.24	0.92	0.00	1.00
Response times (verbs), ms	4542	744.2	101.5	733	374	1284
Motor content (MC)	4565	2.9	1.0	2.8	1.1	6.7
Age-of-acquisition (AoA)	4565	9.9	2.6	10.1	3.1	16.2
Frequency (Zipf)	4471	2.8	0.9	2.8	0.5	6.2
Length—letters (Let)	4565	8.5	2.0	8.0	2.0	17.0
Length—syllables (Syl)	4565	3.4	0.9	3.0	1.0	7.0
Levenshtein's distance (OLD20)	4471	1.78	0.36	1.75	1.0	5.2
Neighborhood—orthographic (OrN)	4471	3.2	2.7	3.0	0.0	25.0
Neighborhood—phonological (PnN)	4471	6.8	5.0	6.0	0.0	50.0
1st-syllable frequency—type (FSFtp)	4565	12,212	14,294	8,016	2.0	46,938
1st-syllable frequency—token (FSFtp)	4565	367,435	431,513	222,227	30.0	1,400,000
Logit prevalence (LogPreval)	4565	2.01	1.47	2.43	−2.67	3.80

Note: Frequency and neighborhood values for 94 verbs were not available in the EsPal database.

(RT < 200 ms or 2 *SD* below the record's mean) or slow (RT > 1500 ms or 2 *SD* above the record's mean) latencies were excluded from the dataset, which represented 4.6% of the total correct responses. The resulting dataset contains 97,072 recognition latencies in ms for 4,542 verbs.

Descriptive statistics

Descriptive statistics and histograms of the distribution of the accuracy and RT means and the variables' scores by items are presented in Table 1 and Figure 1, respectively. Accuracy rate of verbs shows that there are some verbs with no correct responses at all, although others reached the ceiling of performance. The median indicates that most of verbs were recognized correctly by at least 92% of participants. As can be seen in Figure 1, the distribution of accuracy was clearly biased onto correct responses (skewness = −1.534). The distribution of RT values fitted a normal curve but had the typical moderate positive bias of this outcome (skewness = 0.688). Lexical frequency values were transformed into Zipf scores (log10(fpmw)+3; van Heuven et al., 2014), and this was the metric henceforth used in subsequent analyses. Both orthographic Levenshtein's distance (henceforth OLD20) and orthographic and phonological substitution neighborhood sizes were highly skewed (skewness of 2.181, 2.334 and 2.550, respectively) as well as both measures of first syllable frequency (skewness of 1.593 for type frequency and 1.523 for token frequency). Motor content and the number of letters and syllables showed moderate-to-low positive skews (skewness of 0.715, 0.402, and 0.415, respectively). By contrast, AoA had a negative low skew (−0.315), and Zipf showed no skewness (0.018). Regarding the mean and range of the predictor variables, motor content

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

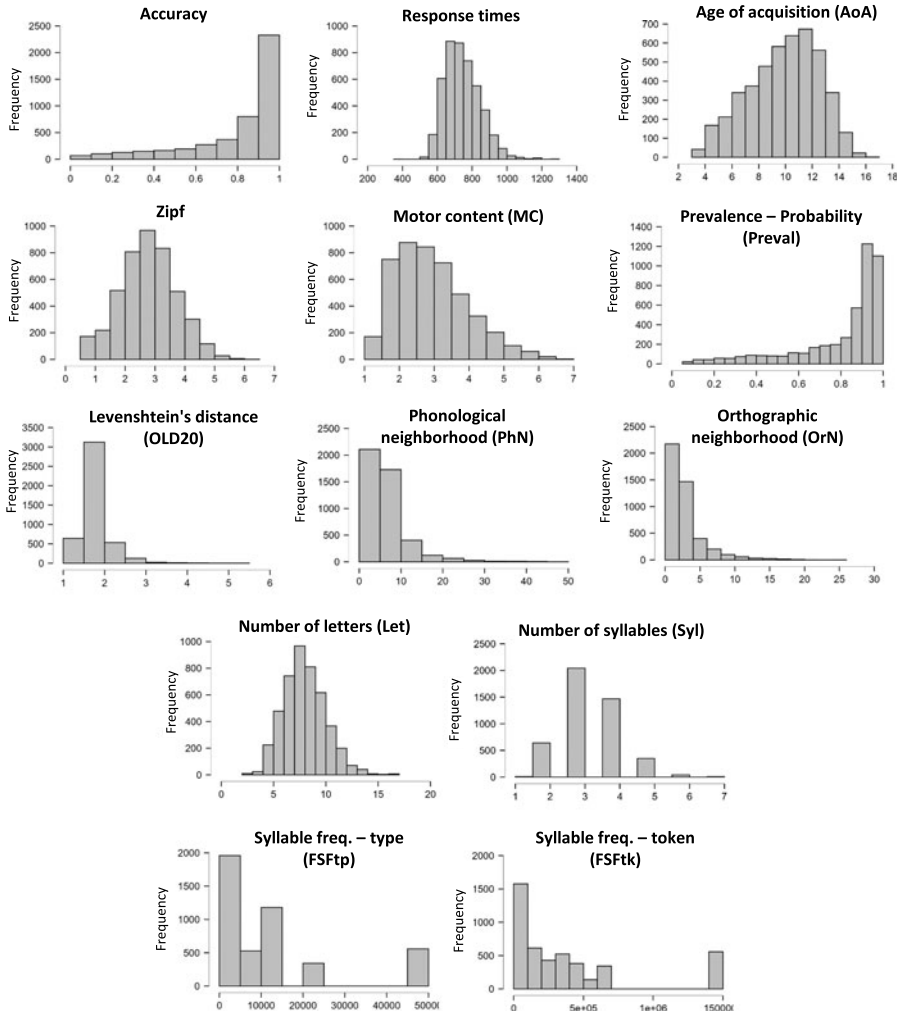


Figure 1. Histograms of each of the variables.

showed scores that cover the entirety of the corresponding 1-to-7 scale with a mean (over 2.9 points) under the mid-point of the scale (i.e. 4 points), AoA scores ranged 13 years, with a mean age near 10 years, and word length and neighborhood measures showed representative means and ranges of the whole Spanish lexicon (i.e., Written and Web corpus of SpaLex: letters, $M = 9.2$, min = 1, max = 23; syllables, $M = 3.8$, min = 1, max = 10; OLD20, $M = 2.3$, min = 1.0, max = 11.3; orthographic neighborhood, $M = 2.2$, min = 0, max = 45; and phonological neighborhood, $M = 4.7$, min = 0, max = 117; see Duchon et al. 2013). Mean and median Zipf frequency values (both over 2.8 points) show that our words are mostly low-to-mid-frequency words as classified according to the intuitive scale suggested by van Heuven et al. (2014) for this variable (i.e., values of

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

1–3 represent low-frequency words, and values of 4–7 represent high-frequency words). However, it must be noted that there is also a considerable number of high-frequency words, with a maximum Zipf value of 6.22. Regarding first-syllable frequency, we find extreme values, with some words having very infrequent first syllable and many words beginning with very frequent syllables (for an extended description of the first-syllable frequency data we refer the reader to San Miguel-Abella et al. 2022). Finally, we used the logit transformation of word prevalence because this has a lower skewness than that of the untransformed prevalence (−0.831 and −1.596, respectively; see Pérez-Sánchez et al., 2026). Overall, the words in the dataset seem to be known enough, low-motor-content, low-to-mid-frequency, late-acquired, with mid-values for length and neighborhood, and beginning with frequent syllables.

Preparation of predictor and dependent variables

Data preparation was targeted to the analyses of words only. First, we transformed both neighborhood and syllable frequency measures to reduce their extreme positive skewness (all > 1.5). After the calculation of the log10(x + 1) of the scores for these measures the skew was 0.465 for OLD, 0.082 for the orthographic neighborhood, −0.370 for the phonological neighborhood, −0.778 for the type syllable frequency, and −0.841 for token syllable frequency. Next, we examined the potential problem of multicollinearity. The problem is that the overlapped information associated with predictors would distort the unique contributions to outcome variance of the given predictors, or it would show unstable effects between different samples (Cohen et al., 2002). High pairwise correlations indicate a potential problem of multicollinearity in a linear model or even in (G)LMM. Therefore, we calculated the pairwise correlations between the psycholinguistic variables, accuracy (in terms of error rate, to facilitate the interpretation), and RT for the word set (see Table 2). We obtained correlations $r \geq .8$, which is a commonly used threshold to suspect a collinearity problem (Cohen et al., 2002), for two pairs and a triad of variables: those for word length (number of letters and of syllables), those for syllable frequency (type and token), and those for neighborhood size (OLD20, orthographic substitution, and phonological substitution). The rest of pairwise correlations between the psycholinguistic variables are $r \leq .62$.

To avoid the potential problem of multicollinearity we selected only one variable of each highly correlated group (pair or triad) as predictor variable in the LMM analyses. We selected in each group the variable with the highest correlation with the corresponding dependent variable, accuracy or latency. Therefore, and for the analysis on accuracy, we selected the number of letters, the (log transformed) phonological neighborhood size, the (log transformed) token first-syllable frequency, and we dropped the number of syllables, the (log transformed) OLD20 and orthographic neighborhood, and the (log transformed) type first-syllable frequency (see correlations in Table 2). Moreover, prevalence, as a derivate of accuracy, was not considered for the analysis on accuracy.

For the analysis on RTs, we selected the number of letters, the (log transformed) phonological neighborhood size, the (log transformed) type first-syllable frequency, and we dropped the number of syllables, the (log transformed) OLD20 and orthographic neighborhood, and the (log transformed) token first-syllable frequency.

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Table 2. Correlation matrix (pair-wise selection)

	1	2	3	4	5	6	7	8	9	10	11	12
1. Error rate	1.000											
2. RT	.620***	1.000										
3. MC	.069***	.039**	1.000									
4. AoA	.549***	.583***	−.066***	1.000								
5. Zipf	−.531***	−.619***	−.100***	−.462***	1.000							
6. Let	−.049***	.378***	−.112***	.235***	−.246***	1.000						
7. Syl	−.025	.366***	−.119***	.221***	−.260***	.858***	1.000					
8. LogOrN	−.156***	−.414***	.040*	−.364***	.375***	−.563***	−.479***	1.000				
9. LogPhN	−.204***	−.481***	.034*	−.372***	.499***	−.616***	−.558***	.858***	1.000			
10. LogOLD20	−.190***	.473***	−.053***	.349***	−.484***	.665***	.598***	−.827***	−.903***	1.000		
11. LogFSFtp	−.005	.086***	−.020	.022	−.008	.108***	.299***	−.025	−.030*	.011	1.000	
12. LogFSFtk	−.029*	.066***	−.040**	.019	.028	.126***	.304***	−.031	−.030*	.012	.988***	1.000
13. LogPreval	−.969***	−.664***	−.078***	−.581***	.578***	.029*	.007	.177***	.224***	−.211***	.000	.026

***$p < .001$, **$p < .01$, *$p < .05$. RT, response times; MC, motor content; AoA, age-of-acquisition; Zipf, logarithmic transformation of word lexical frequency, log10(fpmw) + 3; Let, word length in number of letters; Syl, word length in number of syllables; LogOrN, log transformation of word orthographic neighborhood, log10(OrN + 1); LogPhN, log transformation of word phonological neighborhood, log10(PhN + 1); LogOLD20, log transformation of OLD20 (orthographic Levenshtein's distance), log10(OLD20); LogFSFtp, log transformation of type first-syllable frequency, log10(FSFtp); LogFSFtk, log transformation of token first-syllable frequency, log10(FSFtk); LogPreval, logit transformation of prevalence (predicted difficulty of items).

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

The next step was to *z*-standardize the selected predictors (centering and scaling by their global SD), which facilitates model convergence and is critical for the estimation of interaction effects. Moreover, transforming the predictors into the same scale allows a straightforward comparison of the effects. Finally, following a common practice, we transformed RT to log10(RT) to reduce its skewness in the distribution of latencies, which resulted in 0.274 after transformation ($M = 2.86$, $SD = 0.06$, min = 2.57, max = 3.11). Then, we standardized the log-transformed latencies for a more intuitive interpretation of coefficients (β) and plots, as the predictor variables were also standardized. All trial-level data are available as supplementary materials to this paper for readers who may wish to replicate our analyses or perform new ones.

In addition, to further examine the interpretation of the neighborhood effect, we conducted follow-up analyses on pseudoword trials.[1] For the pseudoword accuracy analysis, timeout responses were excluded, and responses were coded as false alarms (1) when a pseudoword received an erroneous "word" response and as correct rejections, and (0) when it received the correct "pseudoword" response. These data were analyzed with a mixed-effects logistic regression model including OLD20 and pseudoword length (number of letters) as predictors, with random effects for participants and items. For the pseudoword latency analysis, we used responded pseudoword trials only and modeled log-transformed RT as a function of OLD20, response type (false alarm vs. correct rejection), pseudoword length, and the interaction between OLD20 and response type, again including random effects for participants and items.

Model construction

We strived to follow the best practice guidance from the seminal paper by Meteyard and Davies (2020) concerning the use of LMM in psychology. Most of the decisions we made regarding the rationale of the analysis process and results reporting are based on the advice given in that guideline. We agree that reporting results from LMM analyses in a comprehensive and systematic way, such as the one proposed by Meteyard and Davies, may improve the understanding and transparency of the psycholinguistic findings based on that technique. However, due to its magnitude, the process of constructing the best regression models for word recognition accuracy and times is detailed in an appendix to this article (available at https://osf.io/2f9us/). Descriptions of the final models (those that best fit to data) for accuracy and times are presented in Tables 3 and 4, respectively. The fixed effects in these tables show the relative magnitude of the effect (β),[2] the standard error, the 95% confidence intervals, the *z*- or *t*-statistic, and the significance level for the variables and interactions of interest. The significant effects of psycholinguistic variables for recognition accuracy and times are shown in Figures 2 and 3, respectively.

Additional analyses on pseudoword responses

To further characterize the behavior of the pseudoverb materials, false alarms were analyzed with a mixed-effects logistic regression including standardized OLD20 and

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Table 3. Final-model description using accuracy as dependent variable

Fixed effects					
	Coefficients (β)	SE	95% CI (inf/sup)	z	p
Intercept	2.3999	0.0422	2.3171/2.4827	56.815	$<2 \times 10^{-16}$
AoA_z	−0.9212	0.0252	−0.9707/−0.8718	−36.488	$<2 \times 10^{-16}$
Zipf_z	0.9447	0.0273	0.8912/0.9983	34.579	$<2 \times 10^{-16}$
LogPhN_z	−0.0873	0.0327	−0.1514/−0.0232	−2.670	.00759
Let_z	0.3985	0.0287	0.3421/0.4548	13.860	$<2 \times 10-^{16}$
MC_z	−0.0733	0.0212	-0.1150/-0.0315	−3.442	.00058
Zipf_z:Let_z	−0.0062	0.0287	−0.0625/0.0499	−0.218	.82706
Zipf_z:LogPhN_z	0.0061	0.0271	−0.0471/0.0594	0.226	.82109
Zipf_z:AoA_z	0.1764	0.0257	0.1260/0.2268	6.860	7×10^{-12}

Random effects			
	Variance	SD	ICC
Participants (intercept)	0.3091	0.556	.063
Items (intercept)	1.2875	1.135	.263

Model fit				
Marginal R^2	Conditional R^2	Adjusted ICC	Conditional ICC	RMSE
.311	.536	.327	.225	0.304

SE, standard errors; CI, confidence intervals; AoA_z, standardized scores of word age-of-acquisition; Zipf_z, standardized scores of logarithmic transformation of word lexical frequency, log10(fpmw)+3; LogPhN_z, standardized scores of log transformation of word phonological neighborhood, log10(PhN + 1); Let_z, standardized scores of word length in number of letters; MC_z, standardized scores of motor content of verbs; Interaction terms include punctuation mark ":" (colon). Confidence intervals and *p*-values are based on Wald method and asymptotic test, respectively. ICCs are based on the proposals by Nakagawa et al. (2017). Logit model equation, [ACC ~ AoA_z + Zipf_z + LogPhN_z + Let_z + MC_z + Zipf_z:Let_z + Zipf_z:LogPhN_z + Zipf_z:AoA_z + (1 | Participants) + (1 | Items)], where ACC is accuracy (1 = correct answer, 0 = error).

pseudoword length (number of letters) as fixed effects, random intercepts for participants and items, and an uncorrelated by-participant random slope for OLD20. The model showed a significant negative effect of OLD20, such that pseudoverbs with lower OLD20 values were more likely to elicit false alarms, $\beta = -0.625$, $SE = 0.032$, $z = -19.52$, $p < .001$, OR = 0.54, 95% CI [0.50, 0.57]. There was also a significant positive effect of number of characters, $\beta = 0.661$, $SE = 0.030$, $z = 21.89$, $p < .001$, OR = 1.94, 95% CI [1.82, 2.05]. The model converged without singularity, showed no evidence of overdispersion (ratio = 0.71), and yielded Akaike Information Criterion AIC = 71300.77, Bayesian Information Criterion BIC = 71358.96, and logLik = −35644.39.

RTs for responded pseudoword trials were analyzed with a linear mixed-effects model fitted to log-transformed RTs. The model included standardized OLD20, false-alarm status, their interaction, and standardized pseudoword length as fixed effects, with random intercepts for participants and items and an uncorrelated

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Table 4. Final-model description using RT as dependent variable

Fixed effects					
	Coefficients (β)	SE	95% CI (inf/sup)	t	p
Intercept	0.1415	0.0359	0.0711/0.2191	3.941	.0001
LogPreval_z	−0.2826	0.0067	−0.2954/−0.2690	−41.959	2×10^{-16}
Zipf_z	−0.0974	0.0058	−0.1092/−0.0861	−16.577	2×10^{-16}
AoA_z	0.0829	0.0052	0.0721/0.0927	15.749	2×10^{-16}
LogPhN_z	−0.0277	0.0058	−0.0408/−0.0178	−4.739	2×10^{-6}
Let_z	0.1497	0.0053	0.1369/0.1577	28.214	2×10^{-16}
LogFSPType_z	0.0233	0.0039	0.0154/0.0309	5.888	4×10^{-9}
MC_z	0.0118	0.0039	0.0038/0.0193	2.996	.0027
Let_z:MC_z	0.0091	0.0039	0.0014/0.0169	2.314	.0207
Let_z:AoA_z	0.0214	0.0046	0.0110/0.0292	4.595	4×10^{-6}
Zipf_z:AoA_z	−0.0279	0.0053	0.0174/0.0382	5.258	5×10^{-7}
LogPreval_z:Zipf_z	−0.0242	0.0066	−0.0364/−0.0102	−3.635	.0003
LogPreval_z:LogPhN_z	0.0205	0.0055	0.0106/0.0323	3.715	.0002

Random effects			
	Variance	SD	ICC
Participants (Intercept)	0.24071	0.49062	.282
Items (Intercept)	0.03881	0.19700	.045
Blocs (Intercept)	0.00655	0.08093	.008
Residual	0.56883	0.75421	

Model fit				
Marginal R^2	Conditional R^2	Adjusted ICC	Conditional ICC	RMSE
.158	.440	.335	.283	0.743

SE, standard errors; CI, confidence intervals; LogPreval_z, standardized scores of logit prevalence; Zipf_z, standardized scores of logarithmic transformation of word lexical frequency, log10(fpmw)+3; AoA_z, standardized scores of word age-of-acquisition; LogPhN_z, standardized scores of log transformation of word phonological neighborhood, log10(PhN + 1); Let_z, standardized scores of word length in number of letters; MC_z, standardized scores of motor content. Interaction terms include punctuation mark ":" (colon). Confidence intervals are based on Wald method, and t-tests and p-values are based on Satterthwaites's methods for degrees-of-freedom and t-statistics. ICCs are based on the proposals by Nakagawa et al. (2017). Linear model equation, [Log10RT_z ~ LogPreval_z + Zipf_z + AoA_z + LogPhN_z + Let_z + LogFSPType_z + MC_z + MC_z:Let_z + Let_z:AoA_z + Zipf_z:AoA_z + Preval_z:Zipf_z + Preval_z:LogPhN_z + (1 | Participants) + (1 | Items) + (1 | Blocks)], where Log10RT_z = standardized values of logarithmic transformation of response times, log10(RT).

by-participant random slope for OLD20. All fixed effects were significant. RTs showed a negative effect of OLD20, $\beta = -0.026$, $SE = 0.002$, $t = -12.53$, $p < .001$; a negative effect of false-alarm status, $\beta = -0.066$, $SE = 0.002$, $t = -30.50$, $p < .001$; a positive effect of length, $\beta = 0.087$, $SE = 0.002$, $t = 53.43$, $p < .001$; and a significant OLD20 × false-alarm interaction, $\beta = 0.0367$, $SE = 0.002$, $t = 16.95$,

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

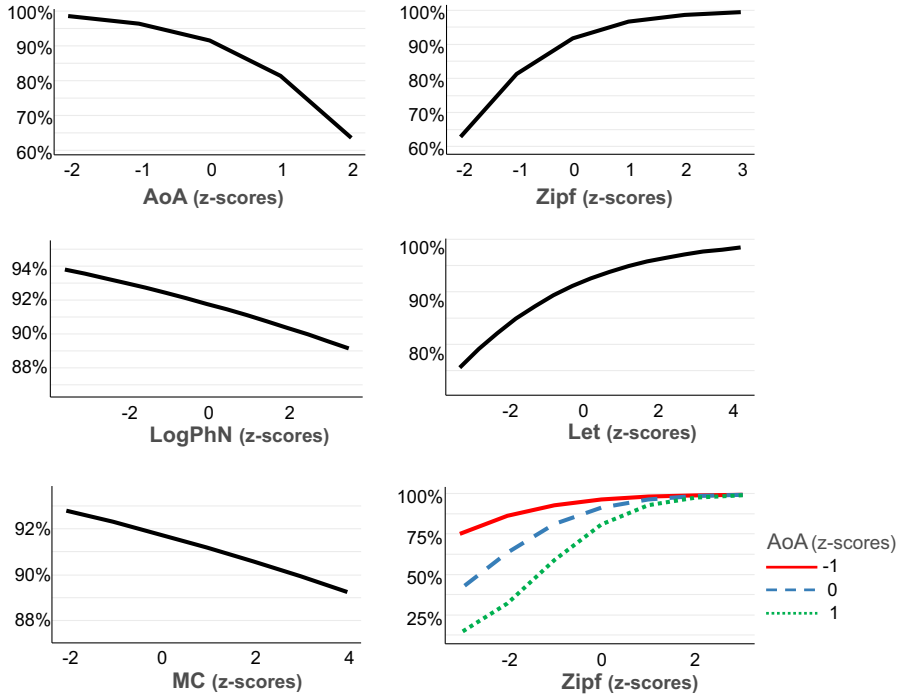


Figure 2. Regression charts of predicted accuracy values by fixed-effect terms (only significant psycholinguistic variables) of the final model.
Note: In the interaction-effects chart, line styles (in color only for the electronic version of the article) represent three values of AoA: the mean, the value one standard deviation above, and the value one standard deviation below the mean (see legend).

$p < .001$. The model converged without singularity and yielded AIC = −20792.35, BIC = −20705.08, and logLik = 10405.18.

Taken together, these analyses indicate that responses to the pseudoverbs were systematically modulated by standard orthographic variables rather than behaving as arbitrary or anomalous items, which supports the adequacy of the final pseudoverb set for use in the lexical decision task.

The data and materials are available in this repository https://osf.io/5cpsh/, and none of the norms or results were preregistered.

The appendix with the regression model construction is available in this repository: https://osf.io/2f9us/ and none of given analyses were preregistered.

Discussion

The main goal of this megastudy is to explore which and to what extent psycholinguistic variables affect visual verb recognition in Spanish with a particular focus on identifying processing patterns that may be specific to the verb category as opposed to lexical processing more generally. The study also examines how these findings relate to those obtained in the verb-naming task reported in SpaVerb-WN, and they

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

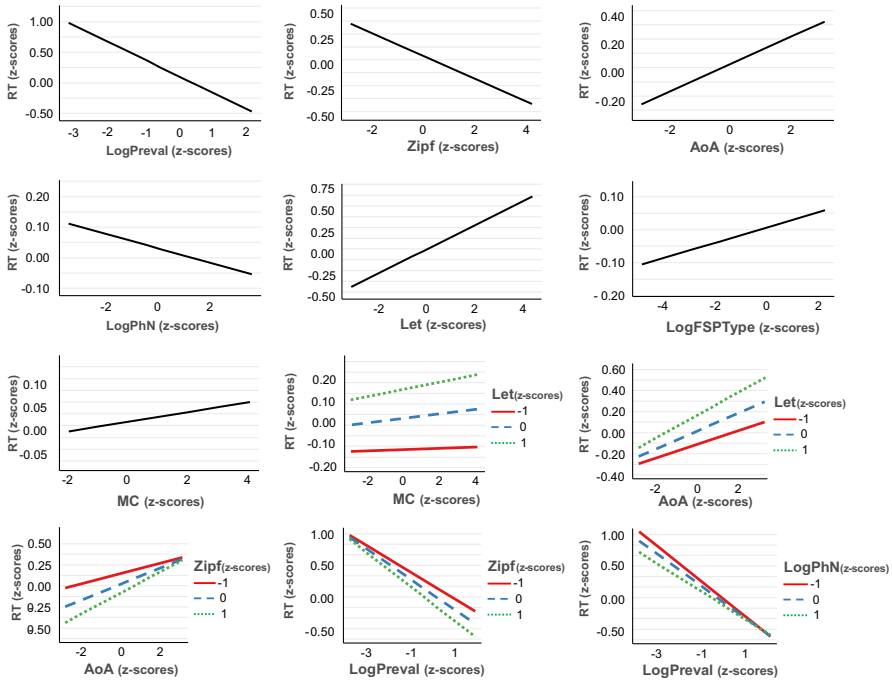


Figure 3. Regression charts of predicted RT (*z* values) by fixed-effect terms of the final model.
Note: In the interaction-effects charts, line styles (and color for the electronic version of the article) represent three values of the corresponding variable (see legend): the mean, the value one standard deviation above, and the value one standard deviation below the mean. RT (*z* values), standardized values of logarithmic transformation of response times, log10(RT).

can be interpreted within broader models of lexical access. In particular, the observed effects of lexical-semantic variables support interactive and embodied accounts, in which semantic information contributes to lexical decision performance, even in transparent orthographies. This contrasts with strictly frequency-based or sublexical models that assume minimal semantic involvement in such tasks.

Word frequency and AoA effects

Word frequency and AoA were, in terms of size effect, the two most relevant variables in predicting response accuracy ($\beta = 0.945$ and -0.921, respectively). Regarding RTs, they were also significant ($\beta = -0.097$ and 0.083, respectively), although in this case behind prevalence and length. Both effects were in the expected direction, i.e., the higher the frequency and lower the AoA of the verb, the faster and more accurate responses. These results are fully consistent with prior evidence, as the effect of these two variables has been consistently observed across different languages (for reviews, see Elsherif et al., 2023; Ghyselinck et al., 2004; Johnston & Barry, 2006; Juhasz, 2005). Regarding the locus of action of these two variables, it has been proposed that both affect lexical access, whereas AoA also influences the access to meanings (see Brysbaert & Ghyselinck, 2006). According to this, the effect

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

of AoA should be more apparent in those tasks that require semantic processing, such as lexical decision, picture naming, or categorization. In reading aloud tasks that do not necessarily require semantic or lexical access, the AoA effect may be more inconsistent and vary across languages. Thus, and in the frame of the dual route theory of reading (e.g., Coltheart et al., 2001), in opaque orthographies such as English, French, or Portuguese, which require lexical-semantic processing for irregular word reading (i.e., thought the lexical route), the AoA effect has been consistently observed (Brysbaert & Cortese, 2010; Cortese & Khanna, 2007; Cortese & Schock, 2013; Soares et al., 2019). By contrast, in transparent orthographies such as Spanish or Italian, where grapheme–phoneme correspondences are highly predictable, semantic or even lexical processing may not be necessary for reading aloud, resulting in the absence of a clear AoA effect (i.e., words can be read through the nonlexical route). Therefore, no evident AoA effect would be observed. However, studies conducted in transparent languages have produced inconsistent findings, with some showing an AoA effect (Bates et al., 2001; Cuetos & Barbón, 2006; Davies et al., 2013; San Miguel-Abella et al., 2022) and others not (Alonso et al., 2016; Alonso et al., 2015; Barca et al., 2002; Burani et al., 2007; Davies et al., 2013). A possible explanation for these mixed results is that the word-naming task may truly involve semantic access depending on the characteristics of words, such that concrete or high-imageability words are read through or with the contribution of the lexical-semantic route. Therefore, the AoA effect would emerge when this kind of words are read (see San Miguel-Abella et al., 2022 for an extended discussion).

Regarding the interaction between word frequency and AoA, the results of our study show that, while analyzing word accuracy, the effect (slope of the regression line) of AoA decreases as word frequency increases, whereas the effect of lexical frequency is more pronounced for late-acquired words ($\beta = 0.176$). However, when we examine RT, the interaction between word frequency and AoA goes in the opposite direction ($\beta = -0.028$). The effect of AoA is more pronounced for high-frequency than for low-frequency words, and the effect of frequency increases as the AoA of words decreases. The effect on accuracy matches with those found in several previous studies in both opaque and transparent languages (Brysbaert et al., 2016; Cortese & Schock, 2013; González-Nosti et al., 2014; San Miguel-Abella et al., 2022). Although our finding for RT is contrary to those reported in most of the previously mentioned studies, it is consistent with that by Brysbaert et al. (2016) in a lexical decision task in Dutch, where the AoA effect was stronger for mid-to-high-frequency words (i.e., Zipf scores from 3 to 5) than for low-frequency words (i.e., Zipf score = 1). This suggests a *multiplicative* effect of word frequency and AoA, where the most frequent words are the most influenced by AoA, and the earliest acquired words are the most influenced by lexical frequency. By contrast, as we noted above, the more consistent interaction found in previous studies between word frequency and AoA seem to be a *mutually compensating* effect; that is, only the less frequent words are influenced by AoA, and only the latest acquired words are influenced by lexical frequency.

One possible explanation for this divergence is that RT and accuracy capture partially different aspects of lexical processing. Although RT is typically more sensitive to the dynamics of lexical access, accuracy may be more strongly influenced by decision-related factors, response uncertainty, or speed–accuracy trade-offs, which could lead to differences in the observed interaction patterns.

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

In addition, the characteristics of the present stimulus set may contribute to this effect. The exclusive use of infinitive verb forms results in a relatively high degree of morphological homogeneity, which may reduce variability at the form level and increase the relative contribution of lexical-semantic variables to performance. Under these conditions, the interplay between frequency and AoA may manifest differently across dependent measures.

Finally, it is also possible that the relationship between AoA and semantic dimensions that are particularly relevant for verbs, such as motor content, plays a role in shaping this interaction. If AoA is partially associated with the richness or type of semantic representations in the verb domain, its effects may vary depending on frequency and on the specific processing demands reflected in RT versus accuracy.

Although these accounts remain tentative, they provide a more specific framework for interpreting the observed pattern and suggest that future work should further examine how lexical, morphological, and semantic factors jointly modulate AoA by word frequency interactions across tasks and measures. However, these effects should be interpreted specifically for verbs, given their semantic complexity and strong links to action representations. Consequently, the observed patterns may not generalize to other lexical categories.

Length effect

A clear length effect was found in the analysis of RT, which tended to be larger as the number of letters increased ($\beta = 0.150$). This effect has been previously observed, although the influence on RT seems to differ across languages. Thus, New et al. (2006) describe a U-shaped length effect in English, with higher TRs for words of short and long words than for words of medium length (five to eight letters). Similar results have shown in languages with intermediate-depth orthography as French (e.g., Ferrand et al., 2010) or Portuguese (Soares et al., 2019). However, in transparent languages, the observed effect, when significant, tends to be linear (Brysbaert et al., 2016). The most widely accepted explanation is regarding the implication in reading of a serial sublexical pathway in charge of graphemes encoding, so that word length (e.g., number of letters) importantly affects the numbers of decoding processes to be executed and, subsequently, the reading time. The problem is that this explanation fits absolutely to what happens in word reading aloud (i.e., word-naming task), but not in word recognition (i.e., lexical decision task) performed by skilled readers, in which the prominent used route is the lexical-semantic route. Therefore, minimal or no effects of word length are expected to occur in lexical decision task. As a matter of fact, in the megastudy conducted by González-Nosti et al. (2014) the main effect of length was not statistically significant in the regression analysis, even though the correlation between word length and mean RT was $r = .360$ ($p < .01$). However, interactions of word length with neighborhood size and AoA were significant. According to the authors, the effect of length might be masked in high-neighborhood or early-acquired words. Not only that, when stimuli were divided into short (three to six letters) and long (7–10 letters), the length effect appeared only in the long word group, which would be reflecting that short words can be recognized with a glance, although long words cannot. This

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

could be because long words need a more complex visual examination (e.g., more than one eye fixation) or a more elaborated morphological processing. In the case of word naming, the length effect could additionally be due to the substantial role of the sublexical route, whose total decoding time is highly influenced by the number of grapheme-to-phoneme conversions required. Both the megastudy by San Miguel-Abella et al. (SpaVerb-WN) and the present study (SpaVerb-LD) revealed a clear inhibitory main effect of stimuli length on RT. Following the explanations above, the length of stimuli could explain the mixed results found. The words selected for the Gonzalez-Nosti et al. megastudy ranged from 3 to 10 letters, with a mean of 6.6, whereas the stimuli selected for the two SpaVerb megastudies are between 2 and 17 letters, with a mean of 8.5. This two-letter difference could explain why the main effect of length is more apparent in both SpaVerb studies. Beyond length differences, a theoretically important distinction is that González-Nosti et al. included words from multiple lexical categories, whereas SpaVerb focuses exclusively on verbs. This categorical focus may modulate the influence of variables such as word length or AoA, as verbs are known to rely more heavily on semantic and sensorimotor information.

In the present study, the aforementioned interaction between AoA and length was also observed for RT ($\beta = 0.021$) and in the same direction described by González-Nosti et al. (2014), that is, the effect of AoA is more accentuated in long words than in short words and the effect of length is larger in late-acquired words than in early-acquired words. Moreover, a small but statistically significant interaction effect between length and motor content was also found in predicting RT ($\beta = 0.009$), with a greater effect of motor content on long words. This study is the first to find such an interaction and complements previous findings (e.g., San Miguel-Abella et al., 2022) about the still novel variable motor content. We propose that both interactions, between AoA and length and between length and motor content, would be specially related to the longer time that long words take to be recognized. Regardless of what is behind the slower processing of long words (see discussion above), the point here is that it may have collateral effects on the variables acting at a semantic level, which in our case are AoA and motor content. As verb processing takes longer, the probability of word meaning involvement in such processing increases, and therefore the effects of semantic variables are also more likely to be large. In sum, the larger effects of AoA and motor content in long verbs arise as an indirect consequence of the greater implication of semantics, which is favored, in turn, by the extra time necessary to process such verbs. These effects are particularly informative for verbs, given their close association with action and sensorimotor representations. The inhibitory effect of motor content, for instance, likely reflects additional processing demands linked to embodied semantic networks that are especially relevant for verbs.

Regarding word accuracy, a main facilitatory effect of the number of letters was also found, with a considerable effect size ($\beta = 0.399$); we observed in our data that the longer words were associated with a lower probability of error. This result fits with previous evidence. Both Aguasvivas et al. (2020) and Yap and Balota (2009) reported similar effects from megastudies in Spanish and English, respectively. However, the reason behind the facilitatory effect of word length on accuracy and its inhibitory effect on RT remains unclear. One possible explanation for this effect of

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

length on accuracy could concern the interaction between word length and frequency (i.e., the word frequency effect decreases as length increases) found in Spanish by Aguasvivas et al. (2020) and in several other studies carried out in different languages (see Barton et al., 2014 for a review). However, we did not find that interaction ($\beta = -0.006$, $p = .827$). One more tentative explanation is related to a potential speed-accuracy automatic trade-off. If longer words require more time to be visually analyzed (i.e., more than one eye fixation), this additional time could contribute collaterally to a sort of *consolidation* effect of the item's information that already entered first in the system (i.e., that from the first eye fixation). Consequently, responses based on that more consolidated information could be more accurate as well. Nevertheless, further research is required to explore this matter thoroughly.

Neighborhood effect

Phonological neighborhood had a slight effect on RT; a main effect of this variable was found ($\beta = -0.028$), but with a smaller size than prevalence, length, frequency, and AoA. The effect was facilitatory so that verbs with a greater number of neighbors were recognized faster. These results are consistent with those found in most of the of previous studies (Aguasvivas et al., 2020; Andrews, 1997; Ferrand et al., 2010; González-Nosti et al., 2014; Yap & Balota, 2009), although mixed results were found by Nemati et al. (2022). The facilitation effect of neighborhood has been attributed to the activation of multiple lexical representations by words with dense neighborhoods, allowing the "word-response" to rely on total activation in the lexicon rather than the activation of a specific representation. Interestingly, the interaction found here between verb prevalence and neighborhood size ($\beta = 0.021$) could be supporting that hypothesis (see Brysbaert et al., 2016, for a similar effect in Dutch). The interaction indicates that the neighborhood effect decreases as prevalence increases, being null at the highest levels of prevalence. Thus, the recognition time for a verb does not depend on its neighborhood size when that verb is clearly known or, in other words, neighborhood size only matters for unknown verb recognition. This can be interpreted as follows. Unknown or less-known words, in terms of prevalence, are responded to in terms of the total activation in the lexicon. In contrast, clearly known words are identified based on the activation of their corresponding specific lexical-semantic representations (see below the discussion about the complementary roles of word prevalence and frequency in lexical decision).

However, the expected interaction between neighborhood and word length was not found (the effect did not contribute significantly to model fit for either accuracy or RT; see Appendix), contrary to findings showing that the effect of neighborhood is mainly observed in short words (Brysbaert et al., 2016; Gonzalez-Nosti et al., 2014). This discrepancy could be attributed to the very low density of neighborhoods in long words. It is also plausible that the greater length of our stimuli in relation to other studies carried out with nouns (see numbers mentioned above) may be behind both the small effect size of neighborhood and the absence of its interaction with other variables. In SpaVerb-WN (San Miguel-Abella et al., 2022), unlike in previous studies in Spanish using the same task (Davies et al., 2013), no

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

main effect of neighborhood or interaction with other psycholinguistic variables was found either, which could support this hypothesis.

Neighborhood was also modest, but statistically significant, in predicting word accuracy ($\beta = -0.087$). The effect was inhibitory because high-neighborhood verbs were recognized with more errors than the low-neighborhood verbs. This result does not fit the facilitatory effect of neighborhood previously found in lexical decision (e.g., Aguasvivas et al., 2020; Ferrand et al., 2010; Yap & Balota, 2009) and does not support the classical explanation given for such effect. By contrast, the present result suggests that high-neighborhood verbs activate a higher number of competitors at the lexical level, and it would hinder the recognition of the target verb and consequently increase the error rate. However, this hypothesis is not endorsed in turn by the facilitatory effect of neighborhood size found on RT here and in previous studies. How can a word be simultaneously facilitated and inhibited in its recognition by the same variable? To address this contradiction, we propose a speculative explanation based on a speed–accuracy trade-off that is compatible with the classical explanation for the facilitatory effect on response speed. If neighborhood size affects such a way that the "word-response" can rely on the total activation in the lexicon rather than on the activation of the individual word representation, this could be viewed as a response bias in the recognition criteria favoring speed over accuracy. In this sense, the "word-response" can be understood as a familiarity-based decision that relies on a global signal of lexical activation, rather than on the full identification of a specific lexical entry, a mechanism akin to the early familiarity check proposed in models such as the EZ Reader framework (Reichle et al., 1998; Reichle et al., 2003). Therefore, the "word-responses" to those items that "sound like words" because they activate many alike words (i.e., high-neighborhood words) trend to be fast but not checked at a lexical level, that is, they have lower RT but are emitted with more inaccuracy.

This interpretation is supported by the additional analyses on pseudoword responses. Specifically, pseudowords with denser orthographic neighborhoods were more likely to elicit false alarms, even after controlling for length. Moreover, false alarms were faster than correct rejections overall, and the effect of neighborhood density differed depending on response type: denser pseudowords were associated with faster false alarms but slower correct rejections. This pattern suggests correct rejections. This pattern suggests that neighborhood density increases the word-likeness of the stimulus and biases participants toward a rapid "word" response when the item is highly similar to existing lexical representations. Therefore, the opposite effects of neighborhood on word RT and word accuracy observed in the main analyses need not be interpreted solely in terms of facilitation versus inhibition at the lexical level, but may also reflect a familiarity-based or criterion-shift component in lexical decision.

Motor content effect

A small but consistent effect of motor content was found, as it appears in both accuracy and RT prediction ($\beta = -0.073$ and 0.012, respectively). This is an inhibitory effect, such that the higher the motor content, the higher the RT and the error rate. Motor content is a measure about the semantic quality of a verb in terms

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

of the amount of displacement or movement of the different parts of the body involved in the action referenced by the given verb (San Miguel-Abella & González-Nosti, 2020). Although there are limited studies comparing these results, as motor content is a relatively new variable that has mainly been tested in patients with motor disorders rather than healthy adults, one relevant study is SpaVerb-WN (San Miguel-Abella et al., 2022). In that study no main effect of motor content was found, but an interaction with AoA was observed. Given that the tasks used in SpaVerb-WN was word naming and in the present one was lexical decision, it can be suggested that differences in the motor content effect may reflect the distinct processing demands of each task. The lexical decision task involves semantic processing and accessing word meaning, which likely influences and improves the response, making it reasonable to expect a significant effect of a semantic variable such as motor content in this context. In contrast, lexical-semantic access is not crucial for word naming, at least in transparent languages where reading can be done through non-lexical routes. Therefore, weak effects of lexical-semantic variables like motor content are expected in word-naming tasks.

It is plausible that motor simulation processes are less engaged in lexical decision than in word naming, where articulation is required. This difference in task demands may partly explain why motor content appears as a main effect in SpaVerb-LD but mainly as an interaction in SpaVerb-WN. Furthermore, the activation of a broader network involving multiple brain areas (i.e., both sensorimotor and perisylvian regions) for high motor-content verbs could increase processing complexity, resulting in longer recognition times and higher error rates.

However, the inhibitory effect of motor content found here does not fit the facilitatory effect shown by an analog variable, the *relative embodiment* (Sidhu et al., 2014, 2016). Relative embodiment refers to the degree to which the meaning of a verb involves the human body. Although content and relative embodiment are not equally defined, their ratings should be positively correlated. Sidhu et al. (2014) concluded that participants activated sensorimotor information from body experience during word reading and that information contributes to the verb-recognizing process, although the authors did not explain exactly how. As mentioned, our motor content effect is inhibitory on recognition latencies, and we think that it could be understood from the so-called *minimal-embodiment* perspective (e.g., Bedny & Caramazza, 2011; Mahon & Caramazza, 2008). This view does not deny the participation of different brain networks in semantic representation, but it argues that networks activate and interact dynamically and differently for action, perception, and language. Word comprehension (spoken or written) may occur primarily in left-hemisphere perisylvian networks, engaging areas associated with modality-independent, conceptual-abstract processing, although activation *can* also spread to other sensorimotor neural circuits. These areas enrich the semantic grounding of words by retrieving specific sensory and motor information, what Mahon and Caramazza (2008) called the "grounding by interaction" framework. Thus, when word reading or recognition is required, the conceptual-abstract network, as part of the linguistic system, could be primarily (required and fast) consulted, although the sensorimotor networks could be done in a subsidiary (supplementary and slow) way.

If high content-motor verbs somehow depict words with scarce or poor conceptual-abstract features in comparison with those of low content-motor verbs

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

(i.e., with extensive and rich conceptual-abstract features), a processing advantage in lexical tasks would emerge for the last, that is, an inhibitory effect of motor content of verbs. Evidence from patients with motor cortex involvement supports this view; these patients show greater difficulty processing action words compared with non-action words, and impaired access to sensorimotor networks may disproportionately affect high motor-content words. Another argument supporting the inhibitory effect of motor content is that the activation of a broader network involving several brain areas (i.e., both the sensorimotor and perisylvian areas) in the case of high motor-content verbs could imply greater processing complexity and thus higher recognition times and errors. This explanation was also proposed by San Miguel-Abella et al. (2022) when discussed the interaction between motor content and AoA that they found in a verb-naming task (i.e., there were an inhibitory effect of motor content in early-acquired words). In the present study, that interaction was not observed and a possible reason is related to the more lexical-semantic nature of the lexical decision task in comparison with word naming. Thus, the motor content effect makes more apparent as a main effect in lexical decision, although it was observed more subtly as a two-way interaction in word naming.

Syllable frequency effect

The present study failed to find effects of syllable frequency on word accuracy (the effect did not contribute significantly to model fit; see Appendix), but an inhibitory effect on RTs was found ($\beta = 0.023$). This result is consistent with those previously reported in Spanish (Carreiras et al., 1993; Carreiras & Perea, 2004; but see Carreiras et al., 2006) and in other languages such as French and German (e.g., Chetail & Mathey, 2009a; Conrad et al., 2006). From the seminal study by Carreiras et al. (1993) the effect of syllable frequency on word recognition has been observed and replicated in several languages (e.g., Carreiras & Perea, 2004; Chetail & Mathey, 2009; Conrad et al., 2006) and using different tasks. Nevertheless, the effect appears in general as facilitatory in word naming (i.e., words with higher syllable frequency are named faster) and as inhibitory in visual lexical decision (i.e., words with higher syllable frequency are identified slower; but see Conrad et al., 2008, and Macizo & Van Petten, 2007). Thus, it is considered that the facilitating effects in word naming occur because in the phonological encoding stage high-frequency syllables are more available and easier to process. In lexical decision, on the contrary, the inhibitory effects of syllable frequency are explained by competition between candidates beginning with the same syllable.

The syllable frequency effect observed in the present study contrasts with the absence of such an effect—whether facilitatory or inhibitory—in the megastudy of verb naming by San Miguel-Abella et al. (2022), SpaVerb-WN. The authors explained the null effect by pointing out the stress homogeneity of Spanish verbs in the forms employed, which always have the stress on the last syllable in infinitive form (e.g., "abanicar," to fan) or on the penultimate one in the pronominal form (e.g., "abanicarse," to fan yourself). As a result, almost all the verbs used in the study had their initial syllables unstressed (only 6 out 4,562 verbs had the stressed syllable in the first position). If the syllable frequency affects only or mainly words with the first stressed syllable, it is logical that the syllable frequency does not affect verb

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

processing. However, we found here that the syllable frequency does affect recognition times of the same set of verbs. Therefore, the explanation provided by San Miguel-Abella et al. for the null effect of syllable frequency cannot be generalized to lexical decision tasks. This, in turn, implies that further research is necessary to clarify whether the association between the syllable effect and the stressed syllable is genuine and, if so, to what extent it is dependent on the task or the stimuli employed.

Word prevalence

Our results on verb prevalence demonstrate that this variable is the most important predictor of recognition times ($\beta = -0.287$), with decreasing times as word prevalence increases. In other words, the knowledge level of verbs predicts the processing speed better than other variables. It is important to underline that other variables, such as word frequency, AoA, and neighborhood size, which are strongly or moderately correlated with word prevalence, also entered in the regression model as significant predictors even when word prevalence was already included. Moreover, word prevalence remained as a significant predictor even after accounting for the other variables examined. This means that word prevalence, on the one hand, cannot be explained by a combination of other well-known psycholinguistic variables, as those mentioned above, and cannot be considered an alternative measure of the lexical frequency.

As far as we know, this is the first study in Spanish that analyses the impact of word prevalence on recognition times (see Brysbaert et al., 2019 for a wide analysis of prevalence as independent and dependent variable). Aguasvivas et al. (2020) addressed word prevalence, which was calculated from SPALEX data (Aguasvivas et al., 2018), as an estimate of vocabulary size, and it was analyzed as a dependent variable in several regression analyses. However, the influence of word prevalence in word-processing times collected in SPALEX has not yet been investigated.

Brysbaert et al. (2016) proposed that word prevalence would be correcting the gaps of lexical corpora based on word occurrence count measures, with those words known by people but that for some reason are rarely talked or written (e.g., words referring to daily life objects, loanwords, and morphologically composed or derived words). The current results fit previous findings regarding the preponderance of word prevalence in predicting word reading and recognition times in Dutch (Brysbaert et al., 2016) and English (Brysbaert et al., 2019). They also support and, to some extent, extend the aforementioned explanation given by Brysbaert et al. about the role of word prevalence in the grammatical category of verbs.

The interactions between word prevalence and other psycholinguistic variables can provide further insights into its role in word processing (see the discussion above regarding the interaction between prevalence and neighborhood). We observed an interaction between word prevalence and word frequency ($\beta = -0.024$), with a practically null frequency effect for low-prevalence words and with a prevalence effect being less pronounced for low-frequency words. Brysbaert et al. (2016, Figure 3) also informed of a similar interaction between these variables (i.e., no frequency effect in low prevalence words and a diminished prevalence effect for the lowest-frequency

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

words), although it was not specifically discussed. We propose that the first side of the interaction (i.e., the almost null frequency effect for low-prevalence words) should be considered as a natural consequence above all other considerations because word frequency cannot affect those unknown words. However, the other side of the interaction (i.e., a less pronounced effect of prevalence in low-frequency words, but considerable in absolute terms) is more complex to explain. We suggest that one explanation of why the prevalence effect persists in low-frequency words can be taken from a suggestion given by Brysbaert et al. (2016; also see Brysbaert et al., 2019) that "not all words require many encounters to be acquired and to be processed rapidly" (p. 452) because there are some words that are acquired easily (better) than others when all of them appear just once (see Borovsky et al., 2012). Brysbaert et al. also discussed that some infrequent words can have a high prevalence because of the semantic depth and richness in which they are incorporated into the semantic system (see also Brysbaert et al., 2019 for an analysis and discussion about whether "word knowledge" implies necessarily semantic information). Therefore, word prevalence effect is related but not absolutely dependent on frequency effect, that is, there can be known words represented in the system, both at lexical and semantic levels, although they were measured with very low or zero frequency scores. By contrast, frequency effect is dependent on prevalence effect, as the former cannot affect unknown (low prevalence) words. If we assume that prevalence effect operates at semantic level and frequency effect does at lexical level, it is reasonable to expect that both effects appear and interact, as observed in our study. However, it should be noted that this distinction is not unequivocal, as word prevalence may also reflect lexical-level familiarity rather than the depth of semantic representations, a possibility already acknowledged in previous work (Brysbaert et al., 2016, 2019). Therefore, the present pattern of results is more appropriately interpreted as converging evidence rather than as a definitive indication of a semantic locus for the prevalence effect.

Finally, we did not find more significant interactions between prevalence and other variables, showing that the main effects of AoA, word length, motor content, and syllable frequency are relatively constant across more or less known words. These results do not replicate the interactions found by Brysbaert et al. (2016) in Dutch, where the prevalence effect was particularly high for late-acquired words, short words, and words with few other alike words. It is also important to note that Brysbaert et al. used mixed lexical categories, whereas our findings come from verbs only. This difference reinforces that lexical category itself can shape how variables like prevalence interact with other psycholinguistic factors. These mixed results reinforce the need for further research on word prevalence and its contribution to word recognition times.

Grammatical class

One objective of the current megastudy was to sum evidence about whether the results obtained with nouns in the lexical decision task are replicated with words of other grammatical categories, particularly with verbs. Previous results in word naming (Davies et al., 2013; San Miguel-Abella et al., 2022) and in picture/action naming (Rodríguez-Ferreiro et al., 2009) provide evidence that such replication may indeed occur. On the other hand, neuropsychological evidence (as mentioned in the

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

introduction) indicates that other variables might be particularly implicated in action/verb processing, such as motor content. Taken together, the present findings highlight that verb processing is not merely a special case of general lexical processing, but instead reflects category-specific dynamics that should be explicitly considered in models of lexical access. These results fit the notion that the processing differences between nouns and verbs (i.e., a processing advantage for nouns) cannot be reduced only to lexical- morphological aspects, but they are related rather to semantic dimensions (e.g., Kauschke & Stenneken, 2008). In this respect, the motor content effects found here on verb recognition latencies can contribute to the knowledge of semantic networks of actions terms (i.e., mainly verbs) in comparison with those of non-action terms (i.e., mainly nouns; see the discussion about motor content above).

Limitations of the present work

Although Spanish serves as the native and official language for millions of speakers worldwide, there exist linguistic and cultural variances among Spanish-speaking populations and communities. In particular, lexical frequency and familiarity may vary across dialects (e.g., certain verbs may be more commonly used in some varieties than in others), which could influence behavioral measures such as RTs. This could potentially impact on the direct extrapolation of the current findings to other populations. However, we believe there's no compelling reason to anticipate divergent results from an equivalent study involving a different Spanish-speaking population, although this assumption depends on the use of dialect-appropriate psycholinguistic norms. In the present study, frequency norms were derived from Peninsular Spanish corpora, and participants were speakers of this same variety, which reduces the likelihood that our results are driven by dialect-specific familiarity. Moreover, all variations of Spanish keep the core characteristic of transparency (i.e., spelling-to-sound correspondences, syllable limits, and accent rules). Other limitation is that behavioral data (accuracy rates and RT) were obtained from a sample of, mostly women, Spanish university students. The presence of educational level and gender biases in our sample is a common occurrence in many psychological studies, and these biases may restrict the generalizability of the results to the broader population. Concerning gender, there is no inherent reason to anticipate significantly different results in a more balanced sample. However, regarding educational level, it is not advisable to directly generalize the present results to populations with differing educational backgrounds, as reading skills and word knowledge are influenced by education.

Conclusions

Under the premise that effects observed in word recognition within one language may not readily apply to others, particularly when they vary in the transparency of their orthographic systems, we undertook a megastudy on lexical decision in Spanish. We explored the role of known and relatively novel sublexical, lexical, and lexical-semantic variables in visual word recognition, taking the data from the

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

megabase by Pérez-Sánchez et al. (2026). Our study utilized the same set of verbs as the word-naming study conducted by San Miguel-Abella et al. (2022), both as a complement and for comparative purposes. Overall, our results fit with those previously obtained with nouns in Spanish in lexical decision. The regression model obtained by LMM on recognition times revealed significant facilitative effects of word frequency and neighborhood size, as well as inhibitory effects of word length, AoA, and syllable frequency. Our study introduced, for the first time in Spanish, the broad impact of word prevalence. This suggests that word prevalence may compensate for certain limitations of word frequency measures and could more likely operate at the semantic, but also at a lexical level. In addition, the inhibitory effects of the motor content of actions provide insights for theories of embodied cognition and may help refine hypotheses regarding difficulties in processing action verbs observed in patients with motor deficits.

Several interaction effects previously reported in the literature were also observed here. The interaction between word prevalence and word frequency suggests that lexical frequency effects depend on whether a word is represented in the mental lexicon, whereas word prevalence captures a broader dimension of word knowledge that can operate even for low-frequency items. The interaction between word prevalence and neighborhood size points to neighborhood effects reflecting a familiarity-based mechanism that operates mainly for low-prevalence words but becomes irrelevant once words are firmly represented in the lexicon. The relationship observed between AoA and length indicates that longer words, due to their increased processing demands, allow greater involvement of semantic information, thereby amplifying the effects of variables such as AoA. Finally, the interaction between AoA and word frequency was paradoxical in relation to most previous findings (i.e., a decreasing frequency effect as AoA increases), suggesting that lexical and semantic factors jointly and dynamically constrain processing, particularly in verbs, although more research is needed to consolidate these findings.

Overall, these findings indicate that, beyond sublexical and lexical processes, semantic access operates during lexical decision tasks in transparent languages. They also emphasize the need for verbs to be investigated in large-scale studies and suggest that future research should incorporate word prevalence and motor content of verbs while designing psycholinguistic experiments or clinical assessments, thereby advancing in their abstract and embodied representational status within lexical-semantic processing.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0142716426100903.

Replication package. All research materials, data, and analysis code are available at:

- Raw data: Pérez-Sánchez et al. (2026) Scientific Data, https://doi.org/10.1038/s41597-025-06339-0
- Data and materials: https://osf.io/5cpsh/
- Appendix with regression model construction: https://osf.io/2f9us/

None of the norms or results were preregistered.

The appendix detailing the regression model construction and the full analysis code used to produce the results reported in the article are both included in this repository.

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Notes

1 We thank an anonymous reviewer for suggesting this analysis.

2 Coefficients (β) from GLMM and LMM can be interpreted as relative effect sizes when predictors are standardized, reflecting the change in the outcome (log-odds for GLMM; SD units for LMM) for a one-SD change in the predictor (Schielzeth, 2010).

References

Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology*, **99**(2), 245–264. https://doi.org/10.1348/000712607X224478

Aguasvivas, J., Carreiras, M., Brysbaert, M., Mandera, P., Keuleers, E., & Duñabeitia, J. A. (2020). How do Spanish speakers read words? Insights from a crowdsourced lexical decision megastudy. *Behavior Research Methods*, **52**(5), 1867–1882. https://doi.org/10.3758/s13428-020-01357-9

Aguasvivas, J. A., Carreiras, M., Brysbaert, M., Mandera, P., Keuleers, E., & Duñabeitia, J. A. (2018). SPALEX: A Spanish lexical decision database from a massive online data collection. *Frontiers in Psychology*, **9**, 2156. https://doi.org/10.3389/fpsyg.2018.02156

Alonso, M., Díez, E., & Fernandez, A. (2016). Subjective age-of-acquisition norms for 4,640 verbs in Spanish. *Behavior Research Methods*, **48**(4), 1337–1342. https://doi.org/10.3758/s13428-015-0675-z

Alonso, M. A., Fernandez, A., & Díez, E. (2015). Subjective age-of-acquisition norms for 7,039 Spanish words. *Behavior Research Methods*, **47**(1), 268–274. https://doi.org/10.3758/s13428-014-0454-2

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin and Review*, **4**(4), 439–461. https://doi.org/10.3758/BF03214334

Ardila, A. (1991). Errors resembling semantic paralexias in Spanish-speaking aphasics. *Brain and Language*, **41**(3), 437–445. https://doi.org/10.1016/0093-934X(91)90165-W

Ardila, A., & Cuetos, F. (2016). Applicability of dual-route reading models to Spanish. *Psicothema*, **28**(1), 71–75. https://doi.org/10.7334/psicothema2015.103

Arévalo, A., Perani, D., Cappa, S. F., Butler, A., Bates, E., & Dronkers, N. (2007). Action and object processing in aphasia: From nouns and verbs to the effect of manipulability. *Brain and Language*, **100**(1), 79–94. https://doi.org/10.1016/j.bandl.2006.06.012

Bak, T. H., O'Donovan, D. G., Xuereb, J. H., Boniface, S., & Hodges, J. R. (2001). Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease-dementia-aphasia syndrome. *Brain: A Journal of Neurology*, **124**(Pt 1), 103–120. https://doi.org/10.1093/brain/124.1.103

Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J. (2012). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Ed.), *Visual word recognition Volume 1: models and methods, orthography and phonology* (pp. 90–115). Psychology Press. https://doi.org/10.4324/9780203107010

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**, 445–459. https://doi.org/10.3758/bf03193014

Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, and Computers*, **34**(3), 424–434. https://doi.org/10.3758/bf03195471

Barton, J. J. S., Hanif, H. M., Eklinder Björnström, L., & Hills, C. (2014). The word-length effect in reading: A review. *Cognitive Neuropsychology*, **31**, 5–6. https://doi.org/10.1080/02643294.2014.895314

Bates, E., Burani, C., D'amico, S., & Barca, L. (2001). Word reading and picture naming in Italian. *Memory and Cognition*, **29**(7), 986–999. https://doi.org/10.3758/BF03195761

Bedny, M., & Caramazza, A. (2011). Perception, action, and word meanings in the human brain: The case from action verbs. *Annals of the New York Academy of Sciences*, **1224**(1), 81–95. https://doi.org/10.1111/j.1749-6632.2011.06013.x

Bocanegra, Y., García, A. M., Lopera, F., Pineda, D., Baena, A., Ospina, P., Alzate, D., Buriticá, O., Moreno, L., Ibáñez, A., & Cuetos, F. (2017). Unspeakable motion: Selective action-verb impairments in

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Parkinson's disease patients without mild cognitive impairment. *Brain and Language*, **168**, 37–46. https://doi.org/10.1016/j.bandl.2017.01.005

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2001). Age-of-acquisition and word frequency in the lexical decision task: Further evidence from the French language. *Cahiers de Psychologie Cognitive*, **20**(6), 401–443. https://doi.org/10.6018/analesps.539801

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, **112**(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Brysbaert, M. (2017). Age of acquisition ratings score better on criterion validity than frequency trajectory or ratings "corrected" for frequency. *Quarterly Journal of Experimental Psychology*, **70**(7), 1129–1139. https://doi.org/10.1080/17470218.2016.1172097

Brysbaert, M. (2022). *Word megastudy data and eye movement corpora available*. 2022. http://crr.ugent.be/programs-data/megastudy-data-available

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology*, **64**(3), 545–559. https://doi.org/10.1080/17470218.2010.503374

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., McCormick, S. F., & Keuleers, E. (2019). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods*, **51**(2), 467–479. https://doi.org/10.3758/s13428-018-1077-9

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, **44**(4), 991–997. https://doi.org/10.3758/s13428-012-0190-4

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, **1**(1), 9. https://doi.org/10.5334/joc.10

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch lexicon project 2. *Journal of Experimental Psychology: Human Perception and Performance*, **42**(3), 441–458. https://doi.org/10.1037/xhp0000159

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica (Amst)*, **104**(2), 215–226. https://doi.org/S0001691800000214 [pii]

Burani, C., Arduino, L. S., & Barca, L. (2007). Frequency, not age of acquisition, affects Italian word naming. *European Journal of Cognitive Psychology*, **19**, 828–866. https://doi.org/10.1080/09541440600847946

Carota, F., Moseley, R., & Pulvermüller, F. (2012). Body-part-specific representations of semantic noun categories. *Journal of Cognitive Neuroscience*, **24**(6), 1492–1509. https://doi.org/10.1162/jocn_a_00219

Carreiras, M., Alvarez, C. J., & Devega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, **32**(6), 766–780. https://doi.org/10.1006/jmla.1993.1038

Carreiras, M., Mechelli, A., & Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Human Brain Mapping*, **27**(12), 963–972. https://doi.org/10.1002/hbm.20236

Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain and Language*, **90**(1–3), 393–400. https://doi.org/10.1016/j.bandl.2003.12.003

Chetail, F., & Mathey, S. (2009a). Activation of syllable units during visual recognition of French words in Grade 2. *Journal of Child Language*, **36**(4), 883–894. https://doi.org/10.1017/S0305000908009197

Chetail, F., & Mathey, S. (2009b). The syllable frequency effect in visual recognition of French words: A study in skilled and beginning readers. *Reading and Writing*, **22**(8), 955–973. https://doi.org/10.1007/s11145-008-9135-9

Chiarello, C., Liu, S., Shears, C., & Kacinik, N. (2002). Differential asymmetries for recognizing nouns and verbs: Where are they? *Neuropsychology*, **16**(1), 35–48. https://doi.org/10.1037/0894-4105.16.1.35

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.) Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203774441

Colombo, L., & Burani, C. (2002). The influence of age of acquisition, root frequency, and context availability in processing nouns and verbs. *Brain and Language*, **81**(1–3), 398–411. https://doi.org/10.1006/brln.2001.2533

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language*, **15**, 245–286.

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

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**, 204–256. https://doi.org/10.1037/0033-295x.108.1.204

Conrad, M., Carreiras, M., & Jacobs, A. M. (2008). Contrasting effects of token and type syllable frequency in lexical decision. *Language and Cognitive Processes*, **23**(2), 296–326. https://doi.org/10.1080/01690960701571570

Conrad, M., Stenneken, P., & Jacobs, A. M. (2006). Associated or dissociated effects of syllable frequency in lexical decision and naming. *Psychonomic Bulletin and Review*, **13**(2), 339–345. https://doi.org/10.3758/BF03193854

Cordier, F., Croizet, J. C., & Rigalleau, F. (2013). Comparing nouns and verbs in a Lexical task. *Journal of Psycholinguistic Research*, **42**(1), 21–35. https://doi.org/10.1007/s10936-012-9202-x

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology*, **60**(8), 1072–1082. https://doi.org/10.1080/17470210701315467

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *Quarterly Journal of Experimental Psychology*, **66**(5), 946–972. https://doi.org/10.1080/17470218.2012.722660

Cuayáhuitl, H. (2004). A syllabification algorithm for Spanish. In A. Gelbukh (Ed.), *Computational linguistics and intelligent text processing* (Lecture Notes in Computer Science, Vol. **2945**, pp. 412–415). Springer. https://doi.org/10.1007/978-3-540-24630-5_49

Cuetos, F., & Barbón, A. (2006). Word naming in Spanish. *The European Journal of Cognitive Psychology*, **18**, 415–436. https://doi.org/10.1080/13594320500165896

Davies, R., Barbón, A., & Cuetos, F. (2013). Lexical and semantic age-of-acquisition effects on word naming in Spanish. *Memory and Cognition*, **41**(2), 297–311. https://doi.org/10.3758/s13421-012-0263-8

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, **45**(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Elsherif, M. M., Preece, E., & Catling, J. C. (2023). Age-of-acquisition effects: A literature review. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **49**(5), 812–847. https://doi.org/10.1037/xlm0001215

Fernandino, L., Conant, L. L., Binder, J. R., Blindauer, K., Hiner, B., Spangler, K., & Desai, R. H. (2013). Where is the action? Action sentence processing in Parkinson's disease. *Neuropsychologia*, **51**(8), 1510–1517. https://doi.org/10.1016/j.neuropsychologia.2013.04.008

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., & Pallier, C. (2010). The French lexicon project: Lexical decision data for 38,840 French words and 38,840 pseudo words. *Behavior Research Methods*, **42**(2), 488–496. https://doi.org/10.3758/BRM.42.2.488

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, **35**(1), 116–124. https://doi.org/10.3758/bf03195503

Gainotti, G. (2004). A metanalysis of impaired and spared naming for different categories of knowledge in patients with a visuo-verbal disconnection. *Neuropsychologia*, **42**(3), 299–319. https://doi.org/10.1016/j.neuropsychologia.2003.08.006

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica*, **115**(1), 43–67. https://doi.org/10.1016/j.actpsy.2003.11.002

Glenberg, A. M. (2015). Few believe the world is flat: How embodiment is changing the scientific understanding of cognition. *Canadian Journal of Experimental Psychology*, **69**(2), 165–171. https://doi.org/10.1037/cep0000056

González-Nosti, M., Barbón, A., Rodríguez-Ferreiro, J., & Cuetos, F. (2014). Effects of the psycholinguistic variables on the lexical decision task in Spanish: A study with 2,765 words. *Behavior Research Methods*, **46**(2), 517–525. https://doi.org/10.3758/s13428-013-0383-5

Herrera, E., Rodríguez-Ferreiro, J., & Cuetos, F. (2012). The effect of motion content in action naming by Parkinson's disease patients. *Cortex*, **48**(7), 900–904. https://doi.org/10.1016/j.cortex.2010.12.007

Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language*, **64**(1), 32–58. https://doi.org/10.1016/j.jml.2010.09.002

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. In *Visual Cognition,* **13**(7–8), 789–845. https://doi.org/10.1080/13506280544000066

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. In *Psychological Bulletin,* **131**(5), 684–712. https://doi.org/10.1037/0033-2909.131.5.684

Juhasz, B. J., Yap, M. J., Raoul, A., & Kaye, M. (2019). A further examination of word frequency and age-of-acquisition effects in English lexical decision task performance: The role of frequency trajectory. *Journal of Experimental Psychology: Learning Memory and Cognition*, **45**(1), 82–96. https://doi.org/10.1037/xlm0000564

Kauschke, C., & Stenneken, P. (2008). Differences in noun and verb processing in lexical decision cannot be attributed to word form and morphological complexity alone. *Journal of Psycholinguistic Research*, **37**(6), 443–452. https://doi.org/10.1007/s10936-008-9073-3

Keuleers, E., & Balota, D. A. (2015). Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments. *Quarterly Journal of Experimental Psychology*, **68**(8), 1457–1468. https://doi.org/10.1080/17470218.2015.1051065

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono-and disyllabic words and nonwords. *Frontiers in Psychology*, **1**, 174. https://doi.org/10.3389/fpsyg.2010.00174

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods,* **44**(1), 287–304. https://doi.org/10.3758/s13428-011-0118-4

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology*, **68**(8), 1665–1692. https://doi.org/10.1080/17470218.2015.1022560

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, **48**(7), 805–825. https://doi.org/10.1016/j.cortex.2011.04.006

Kuperman, V. (2015). Virtual experiments in megastudies: A case study of language and emotion. *Quarterly Journal of Experimental Psychology*, **68**(8), 1693–1710. https://doi.org/10.1080/17470218.2014.989865

Kyröläinen, A. J., Keuleers, E., Mandera, P., & Brysbaert, M. (2021). Affect across adulthood: Evidence from English, Dutch, and Spanish. *Journal of Experimental Psychology: General*, **150**(4), 792–812. https://doi.org/10.1037/xge0000950

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster sensorimotor norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, **52**(3), 1271–1291. https://doi.org/10.3758/s13428-019-01316-z

Macizo, P., & Van Petten, C. (2007). Syllable frequency in lexical decision and naming of English words. *Reading and Writing*, **20**(4), 295–331. https://doi.org/10.1007/s11145-006-9032-z

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, **102**(1–3), 59–70. https://doi.org/10.1016/J.JPHYSPARIS.2008.03.004

Menenti, L., & Burani, C. (2007). What causes the effect of age of acquisition in lexical processing? *Quarterly Journal of Experimental Psychology*, **60**(5), 652–660. https://doi.org/10.1080/17470210601100126

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, **112**, 104092. https://doi.org/10.1016/J.JML.2020.104092

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology*, **91**(2), 167–180. https://doi.org/10.1348/000712600161763

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, **14**(134), 20170213. https://doi.org/10.1098/rsif.2017.0213

Neininger, B., & Pulvermüller, F. (2003). Word-category specific deficits after lesions in the right hemisphere. *Neuropsychologia*, **41**(1), 53–70. https://doi.org/10.1016/S0028-3932(02)00126-4

Nemati, F., Westbury, C., Hollis, G., & Haghbin, H. (2022). The Persian Lexicon Project: Minimized orthographic neighbourhood effects in a dense language. *Journal of Psycholinguistic Research*, **51**(5), 957–979. https://doi.org/10.1007/s10936-022-09863-x

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, **13**(1), 45–52. https://doi.org/10.3758/bf03193811

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

Norris, D. (2013). Models of visual word recognition. In *Trends in Cognitive Sciences*, **17**(10), 517–524. https://doi.org/10.1016/j.tics.2013.08.003

Pérez-Sánchez, M. A., Marín, J., San Miguel-Abella, R., Herrera, E., & González-Nosti, M. (2026). A dataset of word recognition accuracy, times, and prevalence for 4,562 verbs and 4,562 pseudoverbs of Spanish. *Scientific Data*, **13**(1), 25. https://doi.org/10.1038/s41597-025-06339-0

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, **6**, 576–582. https://doi.org/10.1038/nrn1706

Pulvermüller, F. (2013). Semantic embodiment, disembodiment or misembodiment? In search of meaning in modules and neuron circuits. *Brain and Language*, **127**(1), 86–103. https://doi.org/10.1016/J.BANDL.2013.05.015

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, **21**(3), 793–797. https://doi.org/10.1111/j.1460-9568.2005.03900.x

Pulvermüller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex*, **9**(5), 497–506. https://doi.org/10.1093/cercor/9.5.497

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, **105**(1), 125–157. https://doi.org/10.1037/0033-295x.105.1.125

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *The Behavioral and Brain Sciences*, **26**(4), 445–526. https://doi.org/10.1017/s0140525x03000104

Riccardi, N., Yourganov, G., Rorden, C., Fridriksson, J., & Desai, R. H. (2019). Dissociating action and abstract verb comprehension post-stroke. *Cortex*, **120**, 131–146. https://doi.org/10.1016/j.cortex.2019.05.013

Rodríguez-Ferreiro, J., Menéndez, M., Ribacoba, R., & Cuetos, F. (2009). Action naming is impaired in Parkinson disease patients. *Neuropsychologia*, **47**(14), 3271–3274. https://doi.org/10.1016/j.neuropsychologia.2009.07.007

Rüschemeyer, S.-A., Brass, M., & Friederici, A. D. (2007). Comprehending prehending: Neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience*, **19**(5), 855–865. https://doi.org/10.1162/jocn.2007.19.5.855

San Miguel Abella, R. A., & González-Nosti, M. (2020). Motor content norms for 4,565 verbs in Spanish. *Behavior Research Methods*, **52**(2), 447–454. https://doi.org/10.3758/s13428-019-01241-1

San Miguel-Abella, R. S., Pérez-Sánchez, M. Á., Cuetos, F., Marín, J., & González-Nosti, M. (2022). SpaVerb-WN—A megastudy of naming times for 4562 Spanish verbs: Effects of psycholinguistic and motor content variables. *Behavior Research Methods*, **54**(6), 2640–2664. https://doi.org/10.3758/s13428-021-01734-y

Schielzeth, H. (2010), Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, **1**, 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Seidenberg, M. S., & Waters, G. S. (1989). Reading words aloud-a mega study. *Bulletin of the Psychonomic Society*, **27**(6), 489–489.

Sereno, J. A., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory and Cognition*, **25**(4), 425–437. https://doi.org/10.3758/BF03201119

Sidhu, D. M., Heard, A., & Pexman, P. M. (2016). Is more always better for verbs? Semantic richness effects and verb meaning. *Frontiers in Psychology*, **7**, 798. https://doi.org/10.3389/fpsyg.2016.00798

Sidhu, D. M., Kwan, R., Pexman, P. M., & Siakaluk, P. D. (2014). Effects of relative embodiment in lexical and semantic processing of verbs. *Acta Psychologica*, **149**, 32–39. https://doi.org/10.1016/j.actpsy.2014.02.009

Soares, A. P., Lages, A., Silva, A., Comesaña, M., Sousa, I., Pinheiro, A. P., & Perea, M. (2019). Psycholinguistic variables in visual word recognition and pronunciation of European Portuguese words: a

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.

mega-study approach. *Language, Cognition and Neuroscience*, **34**(6), 689–719. https://doi.org/10.1080/23273798.2019.1578395

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, **67**(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Wilson, M. A., Cuetos, F., Davies, R., & Burani, C. (2013). Revisiting age-of-acquisition effects in Spanish visual word recognition: The role of item imageability. *Journal of Experimental Psychology: Learning Memory and Cognition*, **39**(6), 1842–1859. https://doi.org/10.1037/a0033090

REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [online]. Corpus de referencia del español actual. www.rae.es [2008].

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, **60**(4), 502–529. https://doi.org/10.1016/j.jml.2009.02.001

Cite this article: González-Nosti, M., Marín, J., San Miguel-Abella, R., Herrera, E., & Pérez-Sánchez, M. A. (2026). SpaVerb-LD: A megastudy of Spanish verb processing and lexical-semantic dynamics. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716426100903

Downloaded from https://www.cambridge.org/core. 05 Nov 2026 at 07:39:56, subject to the Cambridge Core terms of use.