Age of acquisition and imageability norms for base and morphologically complex words in English and in Spanish

Shakiela K. Davies · Cristina Izura · Rosy Socas · Alberto Dominguez

Abstract The extent to which processing words involves breaking them down into smaller units or morphemes or is the result of an interactive activation of other units, such as meanings, letters, and sounds (e.g., *dis-agree-ment* vs. *disagreement*), is currently under debate. Disentangling morphology from phonology and semantics is often a methodological challenge, because orthogonal manipulations are difficult to achieve (e.g., semantically unrelated words are often phonologically related: *casual–casualty* and, vice versa, *sign–signal*). The present norms provide a morphological classification of 3,263 suffixed derived words from two widely spoken languages: English (2,204 words) and Spanish (1,059 words). Morphologically complex words were sorted into four categories according to the nature of their relationship with the base word: phonologically transparent (*friend–friendly*), phonologically opaque (*child–children*), semantically transparent (*habit–habitual*), and semantically opaque (*event–eventual*). In addition, ratings were gathered for age of acquisition, imageability, and semantic distance (i.e., the extent to which the meaning of the complex derived form could be drawn from the meaning of its base constituents). The norms were completed by adding values for word frequency; word length in number of phonemes, letters, and syllables; lexical similarity, as measured by the number of neighbors; and morphological family size. A series of comparative analyses from the collated ratings for the base and derived words were also carried out. The results are discussed in relation to recent findings.

S. K. Davies · C. Izura (✉)
Department of Psychology, Swansea University of Singleton Park, Swansea SA2 8PP, Wales, UK
e-mail: c.izura@swansea.ac.uk

R. Socas · A. Dominguez
Department of Psychology, Universidad de La Laguna, San Cristobal de La Laguna, Tenerife, Spain

Keywords Age of acquisition · Morphology · Phonology · Semantics · English · Spanish

The phonological, orthographic, morphological, and semantic characteristics of words are central elements in word processing research. This is because finding out how the mind interacts with the structures of language (i.e., its sounds, its letters, its morphemes, and the meaning of its words) has proved to be a successful method of understanding the cognitive basis of language production and comprehension (Brysbaert, Van Wijnendaele, & De Deyne, 2000; Cattell, 1886; Davies, Barbón, & Cuetos, 2013; Duñabeitia, Laka, Perea, & Carreiras, 2009; Frederiksen & Kroll, 1976; Izura, Hernández-Muñoz, & Ellis, 2005; Lavidor & Ellis, 2002; Levelt, 1989; Macizo & Bajo, 2006; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Pérez, 2007).

Morphology, in particular, refers to the compositional structure that meaning has within the words in a language. Thus, words are often composed of smaller meaningful units called morphemes (e.g., *home-work*, *penni-less*, *review-ed*). Morphemes can both stand alone as monomorphemic words (e.g., *truth*) and be bonded to other morphemes as part of polymorphemic words (e.g., *-ful* in *truthful*). Morphologically complex words consist of a core morpheme, called *root*, *base*, or *stem*, and one or several add-on morphemes called *affixes*. In languages such as English and Spanish, affixes can be placed either at the beginning of the word (e.g., *fore-* in *foreword*) as *prefixes*, or at the end of the word (e.g., *-er* in *gardener*) as *suffixes*. Two simple or base words joined together form a *compound* word (e.g., *grandfather*). Generally, one of the words in the compound is the head and the other the modifier (e.g., *blackboard*, where *board* is the head and *black* the modifier). The position of the modifier can vary within and across languages; in English, for example, the modifier tends to be placed first (e.g., *tablespoon*, *wheelchair*), whereas in

Spanish the modifier normally occurs in second place (e.g., *matamoscas* "flyswat", *abrelatas* "can opener").

One conventional classification of morphologically complex words considers, among other things, the modification that the affix *causes* in the meaning of the resulting word. Thus, derivational affixes change the core meaning of the base word and often its lexical or syntactic category (e.g., *walk* is a verb meaning the act of advancing by foot at moderate speed, and *walk-able* is an adjective meaning that something is suitable for walking). In contrast, inflectional affixes keep the meaning and syntactic category of the word constant (e.g., *walk* is a verb meaning the act of advancing by foot at moderate speed, and *walk-ed* is a verb meaning the act of advancing by foot at moderate speed [in the past]).

The question of interest in the field of psycholinguistics is whether the brain relies on the morphological structure of the language when processing linguistic information. The investigation of this issue is challenging, because languages often have exception words that are difficult to classify (e.g., *went* or *fui* ["I went" in Spanish] are verbs that do not use the regular suffix to indicate that the action is occurring in the past). Another difficulty is the heterogeneity found within morphological categories. For example, the semantic change in *talkative*, derived from *talk*, is not equivalent to the change in *motive*, derived from *move*. Both words belong to the same morphological category of derivational words, but somehow the semantic distance between the base and the derived word is larger in *motive* than in *talkative*. Is the brain sensitive to these differences?

Available evidence suggests that such sensitivity exists and that words such as *motive*, *talkative*, *went*, and *worked* are not processed, or even accessed, in the same manner (Plaut & Gonnerman, 2000; Taft & Forster, 1975). Whether the observed differences result from an explicit morphological analysis or whether the cognitive response to morphology is dependent upon orthographic, phonological, and semantic interactive processes is currently under debate.

Taft and Forster (1975) conducted one of the earliest studies looking at morphological processing. They found that English native speakers took longer to recognize non-words stems of prefixed words (e.g., *juvenate*) than pseudostems (e.g., *pertoire*). Longer reaction times were also found when the nonword consisted of a real stem joined with a real prefix (e.g., *dejuvenate*) than when the nonword included a real prefix but an invented stem (e.g., *depertoire*). The explanation offered was that morphemes have a lexical representation that is accessed directly in the case of *juvenate*, and indirectly (i.e., after stripping off the prefix *de-*) in the case of *dejuvenate*. The activation of the existing morphemes causes interference at the time to decide that items such as *juvenate* and *dejuvenate* are nonwords. Taft and Forster proposed a model of word recognition that assumed morphological decomposition at the functional and representational levels.

Subsequent studies investigating the role of morphology in lexical processing have initiated a theoretical controversy in the explanation of the observed morphological effects. Part of this debate relates to whether morphology has an implicit or explicit entity in the cognitive system. Thus, single-mechanism models argue that morphological parsing is embedded in phonological, orthographic, and semantic processing (Devlin, Jamison, Matthews, & Gonnerman, 2004; Gonnerman, Seidenberg, & Andersen, 2007; Plaut & Gonnerman, 2000; Raveh & Rueckl, 2000; Seidenberg & Gonnerman, 2000), whereas dual-mechanism models propose that morphological decomposition occurs but only in words with specific characteristics such as morphological regularity (Burani & Caramazza, 1987; Marslen-Wilson et al., 1994; Pinker, 1991; Prasada & Pinker, 1993; Schreuder & Baayen, 1995).

Marslen-Wilson, Tyler, Waksler, and Older (1994) carried out a cross-modal repetition priming study showing support for the morphological decomposition view. Morphology was examined considering three important factors: the type of affix (inflectional or derivational), the position of the affix (prefix or suffix), and, for the first time in the investigation of morphological processing, the semantic and phonological relationships between the stem in the complex word and the stem in the base word. Semantically transparent words include all inflected words (e.g., *work–worked*) and those derived words whose meaning can be easily guessed from their stems and affixes (e.g., *happy* and *happiness*, *worth* and *worthless*). By contrast, the meanings of semantically opaque words cannot be drawn from the meaning of its components (e.g., the words *department* and *depart* denote very different things). Morphemes have phonological transparency when the stems are phonetically equal in the simple (e.g., *friend*) and complex (e.g., *friendly*) forms of the word, whereas phonological opacity occurs when the phonetics of the stem in the simple (e.g., *sign*) and complex (e.g., *signal*) forms of the word change. Marslen-Wilson et al. (1994) devised a semi-orthogonal manipulation of phonological and semantic relations between primes and targets and found that morphological decomposition was greatly determined by the semantic relationship between the stem and the complex word. Thus, simple words facilitated the processing of complex words if the semantic relation was transparent (e.g., *punish* primed *punishment*), but no priming was found for semantically opaque pairs (*casual* did not prime *casualty*). They argued that semantically transparent words are represented in the brain in a morphologically fragmented manner (i.e., stem+ affix), whereas semantically opaque words (*casualty*) require holistic representation, since accessing the stem (*casual*) offers no help in the comprehension of the word. It is important to note that semantically transparent words include the phonologically transparent (*friendly* from *friend*) and also all of the phonologically opaque (*vanity* from *vain*) words. Marslen-Wilson et al.

(1994) also made a claim about how these representations are accessed, suggesting that whether access occurs in a holistic or affix-stripped manner (as suggested by Taft & Forster, 1975) depends on a combination of factors—such as, for example, whether the complex word is a prefixed or suffixed, whether the word is presented in the auditory or visual modality, and so forth.

A number of other studies have suggested a dual mechanism for morphological parsing, in which words are decomposed at the representational or access level (Pinker & Ullman, 2002; Stanners, Neiser, Hernon, & Hall, 1979) or as in Burani, Salmaso, and Caramazza (1984), who proposed two lexical access procedures—one holistic and one compositional—that are activated in parallel in the recognition process (see also Caramazza, Laudanna, & Romani, 1988; Shreuder & Baayen, 1995).

In contrast, single-mechanism models claim that all words are represented in a similar manner and that the morphological structures of inflected and derived words play no direct role in the way they are processed (Elman et al., 1996; McClelland & Patterson, 2002; Seidenberg & Gonnerman, 2000). Connectionist models of the English past tense, for example, are based on the idea that all kinds of morphologically complex word forms are represented and processed like simple words, through associatively linked orthographic, phonological, and semantic codes, and in terms of activation patterns over units and the weighted connections between them. Therefore, a characteristic of single-mechanism models is an understanding that the semantic and phonological overlap between simple and complex forms is a matter of degree. Morphological effects result from the interaction between orthography, phonology, and semantics; the greater the overlap between orthography, phonology, and semantics, the greater the likelihood of a morphological relationship. Supporting this connectionist perspective, Gonnerman and Plaut (2000) found that priming effects reflected the amount of semantic overlap between word pairs. Single-mechanism models are also supported from the behavior observed in network models. Plaut and Gonnerman (2000) simulated morphological priming in a network learning either a morphologically rich language (e.g., Hebrew) or a morphologically weaker language (e.g., English). The English-trained network exhibited priming only for those pairs that were semantically related, whereas the Hebrew-trained network showed priming for those items that were morphologically but not semantically related. These results support other findings (Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000; Marslen-Wilson et al., 1994) and suggest that morphological effects in the absence of semantic overlap can be explained within the connectionist framework. This type of morphological priming in the absence of semantic overlap would only be observable in languages with rich morphology in which the ubiquitous morphological structure dominates the internal

representations of the network. It is important to note that single-mechanism models do consider morphology an important level of analysis; their fundamental difference from dual-mechanism models resides in the fact that the connectionists approach does not conceive morphology as a process independent from phonology, orthography, and semantics.

Notably, the controversy between single- and dual-mechanism models is not only theoretical, but also methodological. Finding morphologically related words that do not overlap phonologically and/or semantically is a research challenge. Words such as *beauty*, *beautiful*, and *beautifully* are all part of the same morphological family, but in addition they also share a great part of their orthography, phonology, and semantics. The present study is an attempt to ease these methodological difficulties by providing norms for morphologically complex words that are phonologically transparent, phonologically opaque, semantically transparent, and semantically opaque in two languages with slightly different morphological structures: English and Spanish. Both languages are rich in relatively different aspects of their morphological compositions. Having said that, inflectional morphologies in most, but not all, languages mark relations such as number, gender, tense, and so forth, in similar manners (e.g., adding "-s/-es" to generate the plural form), providing the possibility of cross-linguistic comparisons (Cutler, Hawkins, & Gillian, 1985; Ramirez, Chen, Geva, & Yang, 2011). In terms of derivational morphology, however, Spanish as a Romance language has a greater abundance of affixed words, whereas English, as a Germanic language, makes more productive use of compounding as a method of word formation (Piera, 1995).

The English and Spanish languages were selected because they are the second and third most widely spoken languages in the world (Weber, 1997). If, in addition, we consider that it is more common to speak two than to speak one language, the number of English–Spanish bilinguals is likely to be high. These norms, therefore, aim to be a useful source of material for research based on monolingual and bilingual speakers. Considering the characteristics of words in the two languages of the bilingual speaker is important, because it has been shown that the processing of words in one language is affected by the orthographic, phonological, and morphological characteristics of the words in the other language (Dijkstra, Moscoso del Prado Martín, Schulpen, Schreuder, & Baayen, 2005; Van Hell & Dijkstra, 2002; Van Heuven et al. 1998).

A final consideration when developing these norms was to include some of the key factors known to affect the ways in which simple and complex words are processed and/or represented in the mental lexicon (see Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004, for a review). The factors included here are age of acquisition, imageability, word frequency, word length (i.e., numbers of letters and phonemes and syllables), and lexical similarity, as measured by the number of orthographic neighbors and morphological family size.

A brief overview of the main findings in relation to these characteristics is provided below.

Age of acquisition

Age of acquisition (AoA) refers to the order in which words, faces, objects, and any other materials for which numerous examples exist are learned. The AoA effect typically emerges whenever stimuli have to be learned over a period of time, in a consecutive and cumulative manner (see Johnston & Barry, 2006, for a review), indicating that the AoA effect might be a key characteristic of learning. AoA effects have also been found in first and in second languages learned after childhood; for this reason, some studies have referred to AoA as to *order of acquisition* (see Izura & Ellis, 2004; Izura et al., 2011; Stewart & Ellis, 2008). Order of learning has a significant influence on the time and precision at which information is processed, as well as being a powerful determinant of the information that is more likely to be lost after brain injury. The effect is such that material learned first is processed quickly and accurately and is unlikely to be lost. The influence of AoA has been shown in both lexical and nonlexical tasks (Bonin, Chalard, Méot, & Barry, 2006; Brysbaert & Cortese, 2011; Brysbaert et al., 2000; Catling, Dent, & Williamson, 2008; Holmes, Fitch, & Ellis, 2006; J. Monaghan & Ellis, 2002; Richards & Ellis, 2009); in many languages (Alija & Cuetos, 2006; Ferrand et al., 2011; Izura & Ellis, 2002; Liu, Hao, Shu, Tan, & Weekes, 2008; Menenti & Burani, 2007; Raman, 2006; Wilson, Cuetos, Davies, & Burani, 2013; Wilson, Ellis, & Burani, 2012); in old adults, young adults, and children (Cuetos, Samartino, & Ellis, 2012; Morrison & Ellis 1995); in bilingual and monolingual speakers (Assink, van Well, & Knuijt, 2003; Hirsh, Morrison, Gaset, & Carnicer, 2003; Izura & Ellis, 2004); and in studies using behavioral and neural responses (Cuetos, Barbón, Urrutia, & Domínguez, 2009; Ellis, Burani, Izura, Bromiley, & Venneri, 2006; Juhasz & Rayner, 2006; Pérez, 2007; Weekes et al. 2008). In sum, the evidence shows that AoA effect is a pervasive property of the organization and function of the cognitive system, applicable to all information learned in a cumulative and interleaved manner.

AoA is commonly measured by asking groups of individuals to estimate the age at which they believe they learned a list of words. These estimations have been shown to correlate highly with objective measures of AoA (Carroll & White, 1973; Gilhooly & Gilhooly, 1980; Pérez, 2007).

As we stated above, the reality of an AoA effect in word processing is supported by ample evidence. However, few if any of the investigations of morphological processing have taken AoA into account. We believe this practice has to change. One study hinting at the relation between AoA and morphology is that of Kuperman, Stadthagen-González, and

Brysbaert (2012). They showed that the AoA of the base word is a significant predictor of the time in which the inflected form of a word is recognized (i.e., the AoA of *play* influences the time in which *played* is recognized). One possible explanation for such an effect is that the AoA of the inflected word form coincides to a large extent with that of the base, as was suggested by Kuperman et al. (2012). Another possibility is that processing words requires morphological decomposition. In this case, the AoA of the stem or base word could have an effect on how quickly the word is decomposed and/or assembled together thereafter. In the present study, a mean comparison of the ratings provided for the base and morphologically complex words was run in an attempt to shed some light on this issue. Thus, if the AoA of base words can be applied to the inflected word forms, no significant differences would be observable between the ratings of base and complex word forms.

Imageability

Imageability refers to the ease with which the meaning of a word evokes a mental image. This factor was developed in the 1960s as a potential explanation of the faster and more accurate processing of imageable than of abstract words. Subsequently, Paivio (1971; Clark & Paivio, 1991) formulated the *dual-code hypothesis*, in which he proposed that words have two potential codes of representation: a verbal and a visual code. Imageable words have an advantage because they enjoy visual and verbal representations, whereas unimageable words are only represented verbally. Evidence shows that highly imageable words are recognized and memorized better in tasks of lexical decision and cued and free recall (Balota et al., 2004; Kennet, McGuire, Willis, & Shaie, 2000). High-imageability words are also less prone to naming errors by patients with phonological impairment (Hirsh & Ellis, 1994; Tree, Perfect, Hirsh, & Copstick, 2001).

Of relevance to the present study is Reilly and Kean's (2007) novel account of the differences between imageable and abstract words. They argued that the relationship between form and meaning is not completely arbitrary or orthogonal, but is interactive, and that the observed differences between imageable and abstract words can be explained because they are different not only at the semantic but at the formal level. Reilly and Kean analyzed a corpus of 2,023 English nouns and found that low-imageability or abstract words are more complex morphologically (i.e., formed by multiple affixes), as well as longer, etymologically more diverse, and more dissimilar to other lexical entries. They argued that these findings reflect properties of the language, showing, for example, that abstract nouns are most commonly created through the affixation of imageable stems (e.g., *man*, *manliness*). In the present study, Reilly and Kean's proposal was tested by comparing the mean ratings for base and morphologically complex

words. According to Reilly and Kean's account, base words should show significant higher imageability values than complex word forms.

Reilly and Kean (2007) also hypothesized a potential developmental explanation for the differences between abstract and imageable words. The claim was supported by the fact that children's vocabularies are populated by an abundance of imageable and short words—possibly, they argued, to reduce cognitive load and facilitate fast semantic mapping. Reilly and Kean's propositions were tested in the present study by running a mean comparison between the imageability ratings given to base words and morphologically complex words. According to Reilly and Kean, base words should come up with significantly higher imageability ratings than complex words. The developmental hypothesis from these authors also suggests that base words should on average be acquired earlier than complex word forms.

Imageability is not a factor regularly considered in studies investigating morphological factors. However, its influences in cued recall, free recall, lexical decision, and naming should be sufficient for researchers not to dismiss it as a potential intervening factor. In addition, in a recent study looking at the storage of inflected word forms, Prado and Ullman (2009) showed that the imageability of the stem (e.g., *walk*) was a significant predictor of the reaction times and acceptability ratings of verbs in sentence completion tasks. They argued that imageability could be a useful tool for investigating the storage and representation of simple and complex words.

Lexical similarity

A common way of measuring lexical similarity is by counting the number of words that can be formed by changing one letter from a given word. Following this rule, *beach*, for example, has seven neighbors: *beech*, *belch*, *bench*, *leach*, *peach*, *reach*, and *teach*. This factor is commonly known as the *number of orthographic neighbors* (*N*), and it was originally proposed by Coltheart, Davelaar, Jonasson, and Besner (1977). A common finding is that words with high numbers of orthographic neighbors are named and recognized faster than words with low numbers of orthographic neighbors (Andrews, 1989, 1992; Mathey, 2001; Perea & Rosa, 2000; Sears, Hino, & Lupker, 1999).

The investigation of morphological decomposition within visual word recognition has often employed priming as the paradigm of study. In this context, the consideration of orthographic similarity between the prime and target has been of paramount importance (Duñabeitia, Perea, & Carreiras, 2007; Marslen-Wilson et al., 1994; Rastle et al. 2004).

Morphologically complex words tend to be long, and as such we do not expect to find large *N* differences between different types of complex words (i.e., whether phonologically transparent, phonologically opaque, semantically transparent, or semantically opaque). However, large *N* differences can be observed in base words. If morphological decomposition is assumed to occur at early stages of visual word processing (Christianson, Johnson, & Rayner, 2005; Rastle & Davis, 2003; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Rastle et al., 2004; Taft, 1994), differences in lexical similarity should be taken into account.

Orthographic and phonological length

Word length, measured by its visual (number of letters) or auditory (numbers of phonemes and syllables) characteristics, shows a positive correlation with word naming and recognition times (Balota et al., 2004; Hudson & Bergman, 1985). A number of studies have revealed a progressive time cost in naming and recognition times as the length of a word increases (Balota et al., 2004; Frederiksen & Kroll, 1976; Juhasz & Rayner, 2003; Spieler & Balota, 1997). However, a recent study examining recognition times of 33,006 English words with letter lengths ranging from three to 13 showed a curvilinear effect of length, with facilitation for words three to five letters long, null effects for words five to eight letters long, and a time cost for words eight to 13 letters long (New, Ferrand, Pallier, & Brysbaert, 2006). This curvilinear relationship between letter length and recognition times can have important implications for theories of morphological processing. Thus, the overall length of a complex word should not matter much for those theories claiming that morphologically transparent or regular words are always assembled (Marslen-Wilson & Tyler, 1998; Taft, 1979; Tyler et al., 2002). Nevertheless, the chances of morphological decomposition may also rely on the lengths of the constituent morphemes (see, e.g., Kuperman, Bertram, & Baayen, 2010).

Word frequency

The number of times an individual encounters a word in the spoken or visual modality is perhaps one of the first and best-studied factors in word recognition research (Cattell, 1886). There is little doubt, nowadays, that word frequency is a powerful determinant of word processing, with an advantage for high-frequency words. Its effect has been shown in word identification, word naming, object naming, recall, translation, categorization, and learning, among other processes (Connine, Mullenix, Shernoff, & Yelen, 1990; Criss, Aue, & Smith, 2011; MacLeod & Kampe, 1996; Murray & Forster, 2004; Oldfield & Wingfield, 1965; Strain, Patterson, & Seidenberg, 1995; Whaley, 1978; Yonelinas, 2002). Through the years of research on word processing, a number of factors have been put forward as potential confounds of word frequency (e.g., AoA, concreteness, contextual diversity, imageability, word length, number of orthographic and/or phonological

neighbors, etc.). However, the frequency of the word has remained as a variable that explains, over and above the confounding factors, a significant proportion of the variance associated to the precision and time associated with processing a word.

The robustness and widespread influence of word frequency suggests that frequency shapes the word's representation in memory, and as such, most models of word processing incorporate frequency in their architectures (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 2004; P. Monaghan & Ellis, 2010).

A number of studies of morphological processing have manipulated frequency, considering that if it is an intrinsic part of the lexical representation, it is a safe factor to use when testing theories of lexical access, storage, and/or retrieval. Taft (2004), for example, used a common argument in his investigation of the recognition of complex words. He manipulated stem frequency (i.e., the base frequency of the stem and its inflected forms; e.g., *walk*, *walking*, *walked*, *walks*) and surface frequency (i.e., the frequency of the word itself; e.g., *walking*). The idea is simple: If the process of word identification is fragmented into stem plus affix, then complex words with high-frequency stems will be identified faster than complex words with low-frequency stems, assuming a control of the overall frequency. Similar design ideas have been exercised in a number of studies of morphological processing (Burani & Caramazza, 1987; Domínguez, Seguí, & Cuetos, 2002; Taft, 2004; Vannest, Newport, Newman, & Bavelier, 2011).

Here we present the surface frequency of the base word (i.e., its own singular frequency) and of the corresponding complex word(s). Surface frequency has been shown to affect reaction times to simple and complex derived words in lexical processing (Baayen, Dijkstra, & Schreuder, 1997; Ferrand et al., 2011; Taft, 1979). In addition, the surface frequency of the exact orthographic configuration of a word has been shown to affect language acquisition and grammatical development. Distributional frequencies, for example, provide important grammatical cues, such as word class, that are picked up by children when learning the language (Diessel, 2007). High-frequency words occurring in particular clusters (e.g., *in the _, I am _, a big _, has been _*) have a tendency to be learned earlier than low-frequency words, having an impact on the emergence of the structure of language. The selection of surface frequency for the present norms is attributable to its importance not only in language processing, but also in acquisition and development.

Surface frequency values gathered from subtitles are the values that we opted for, because subtitle frequencies are available for the Spanish and English languages and are also thought to be more representative of the language in use than are printed frequencies (Cuetos, González-Nosti, Barbón, & Brysbaert, 2011; Dimitropoulou, Duñabeitia, Avilés, Corral, & Carreiras, 2010).

Morphological family size

The morphological family size of a word is the number of different complex word types in which a base word occurs as a constituent (De Jong, 2002). Morphological family size has a significant effect on word recognition, with words from large morphological families being identified faster than those from small morphological families (Bertram, Baayen, & Schreuder, 2002; Mulder, Dijkstra, Schreuder, & Baayen, 2014; Schreuder & Baayen, 1997).

The role of morphological family size is gaining prominence in relation to other fundamental lexico-semantic variables. Across numerous languages, quicker naming times have been associated with words that come from larger orthographic and semantic families (Schreuder & Baayen, 1997; Taft, 1979), and thus with larger morphological family size. In addition, highly frequent morphologically simple words tend to have a larger number of morphological families than low-frequent morphologically simple words (Schreuder & Baayen, 1997). As a consequence, words with higher-frequency morphological families are processed quicker than those words that come from lower-frequency morphological families. This has been evident in morphologically simple (e.g., Baayen, Feldman, & Schreuder, 2006) and morphologically complex (Kuperman et al., 2010) words.

It has been argued that morphological family size and frequency may well be semantic in nature, due to the semantic characteristics between morphological family members (De Jong, 2002; De Jong, Schreuder, & Baayen, 2000). Schreuder and Baayen (1997) showed that the morphological family size effect is indeed, in great part driven by semantics, since only the number of semantically transparent relatives showed a significant influence on their identification task.

Interestingly, morphological family size appears to influence words acquired at different points in time. For example, Henry and Kuperman (2013) examined AoA norms for morphological families and their shared morphemes. Henry and Kuperman observed that those words with larger morphological family sizes had earlier-acquired AoA ratings than those words with fewer morphological family members. Henry and Kuperman explained their findings in relation to the fact that the architecture or the size and development of the mental lexicon is contingent not only on how *early* a word is acquired (AoA of the target morpheme), but also on its family size (Steyvers & Tenenbaum, 2005). Notably, morphological family size has also been shown to remain a significant predictor in other studies, even when AoA has been controlled (De Jong, 2002).

The effect of morphological family size has also been observed in younger readers (Perdijk, Schreuder, Baayen, & Verhoeven, 2012), which one would expect not to find (Schreuder & Baayen, 1997), on the premise that language development for much younger readers is still growing

tentatively. However, Perdijk et al. interpreted the observation of a morphological family size effect emerging in younger readers to be due to the increased use of phonological recoding in the way that young readers process vocabulary.

In order to facilitate future research related to morphological family size, two measures of morphological size are provided here: (1) total morphological family size and (2) the morphological family size of derived words that are semantically related to the base form.

Norms for the present database

The growing number of investigations on the morphological impact of complex words on mental processes has prompted the need to gather norms for families of words that are morphologically related. Here we present norms for 2,204 morphologically related words in English and 1,059 morphologically related words in Spanish. The words were selected according to the phonological or semantic relationships established between the base and the morphologically complex word. The defining criterion for *phonological transparency* was whether or not the stem of the complex word preserved the same phonetic shape as the base word (Marslen-Wilson et al., 1994). Thus, complex words were phonologically transparent if their stem was pronounced the same way in the complex and simple versions of the word (e.g., *friend–friendly* in English; *comer–comedor* "eat–dining room" in Spanish). Phonological opacity occurred when the pronunciation of the base word was different from the pronunciation of the stem in the complex word (e.g., *sign–signal* in English; *cuerpo–corporal* "body–corporal" in Spanish). It is important to note that differences in orthographic transparency and grammatical class (e.g., filtering for nouns and verbs) were not taken into account in the formation and analysis of the sets of words comprising these norms. Furthermore, words were classed as phonologically transparent even if there was a change in the syllable status of a base word relative to its derived complex form. Overall, words were also classed as phonologically opaque if there was change in a vowel or consonant between the base and the derived form (e.g., *complete–completion*) (Marslen-Wilson et al., 1994). The defining criterion for s*emantic transparency* was whether the meaning of the complex word could be guessed via the meanings of the base word and the affix. Thus, semantically transparent words were those easily understood via the comprehension of the base word and the affix (e.g., *bake–baker* in English; *vista–visión* "view–vision" in Spanish). Semantically opaque words were etymologically related to the base word but had meanings that could not be easily guessed from the base word (e.g., *wit–witness*, in English; *calor–caloría* "heat–calorie" in Spanish).

Words were presented in four blocks (two of English words [A and B] and two of Spanish words [C and D]). Each block comprised morphologically related words grouped in pairs or trios. Block A comprised 936 English words. These were grouped in 312 trios of morphologically related words in which the complex forms were phonologically transparent or opaque in their relation to the base word. Block B comprised 1,268 English words, grouped into two pairs of words morphologically related with a semantic relation of transparency (317 pairs) or opacity (317 pairs). Block C consisted of 480 Spanish words. Half of these words were base words and half were morphologically derived words; from these, 120 were phonologically transparent in relation to their base word (e.g., *comer–comedor*, "eat–dining room"), and 120 were phonologically opaque in relation to their base word (e.g., *cuerpo–corporal* "body–corporal"). Block D consisted of 579 Spanish words, grouped in trios of 193 words each. The trios were formed by a base word and two morphologically related words, one semantically transparent in relation to the base word (e.g., *amar–amante* "to love–lover") and one semantically opaque in relation to the same base word (e.g., *invierno–invernadero* "winter–greenhouse").

In addition to the morphological relationship between the complex and simple forms of the word, norms were gathered for the following variables: AoA, imageability, and degree of semantic opacity. Statistical comparisons between the ratings provided for base and complex word forms on AoA and imageability were run to test the proposals made by Reilly and Kean (2007) and Kuperman et al. (2012). Values for word frequency, number of phonemes, number of letters, number of syllables, number of orthographic neighbors, and morphological family size for all of the words considered in the study (i.e., simple and complex words) were also included, to complete the norms.

Method

Participants

A total of 277 English native speakers, 61 males and 216 females, and 256 Spanish native speakers, 49 males and 207 females, participated in the compilation of these norms. Each of the factors to be estimated—AoA, imageability, and semantic similarity—was rated by two different groups of volunteers, one by English native speakers and one by Spanish native speakers. The participants had a mean age of 22 years (range 18 to 47 years) and completed the questionnaires online. They all had normal or corrected-to-normal vision.

Materials

English words (2,204 words) The list of suffix-derived words that were morphologically related and had a relation with the stem of phonological transparency or opacity (*n*= 936 for

English and *n* = 244 for Spanish) were selected from the General Service List (GSL) database of morphologically related English words (West, 1953) and from the Dictionary of the Spanish Real Academy (Real Academia Española, 2001). The selection of English and Spanish words was based on the following criteria: Morphological families had a base word that could form a phonologically transparent and a phonologically opaque word. Complex words were phonologically transparent when their stem was phonetically the same in the simple form of the word (e.g., *friend–friendly*; *comer–comedor* "eat–dining room"). Complex words were phonologically opaque if their stem was phonetically different from that in the simple word (e.g., *sign–signal*; *diez–decena* "ten–tenth").

On the basis of these criteria, all of the English words from the GSL database (West, 1953) were initially coded as phonologically transparent and phonologically opaque, irrespective of their semantic transparency. The nature of the GSL, a database comprising 2,000 base words and their morphological families, allowed for the selection of 312 base words that had a corresponding phonologically transparent derived word (*n* = 312) and also had a phonologically opaque derived word (*n* = 312). Thus, a base word such as *explain* was selected because we could find a phonologically transparent derived word, *explaining*, and a phonologically opaque derived word, *explanatory*, within the same morphological family. The total of phonologically related English words was 936 (see Table 1 for examples). To our knowledge, there is not a database of word families in Spanish; therefore, the selected Spanish words were two lists of word pairs: one list formed with pairs comprising base and morphologically transparent words (119 pairs, e.g., *mora–morado* "blackberry–purple"), and the other list comprising a different set of base and morphologically opaque words (119 pairs, e.g., *boca–bucal* "mouth–oral"). The degree of semantic transparency was not considered when selecting derived words phonologically related to the base word.

The set of morphologically related words for English, with a relation with the stem of semantic transparency or opacity (*n* = 1,252), was selected from the CELEX database (Baayen,

Piepenbrock, & van Rijn, 1993; Lavric, Rastle, & Clapp, 2011; Marlsen-Wilson, Bozic, & Randall, 2008; Morris, Frank, Grainger, & Holcomb, 2007; Rastle, Davis, & New, 2004) and the Concise Oxford Dictionary (Fowler & Fowler, 1982); the words for Spanish were selected from the Dictionary of the Spanish Real Academy (Real Academia Española, 2001).

Finding trios of morphologically related English words whose composition was base word, semantically transparent related word, and semantically opaque word proved to be difficult. Therefore, words were grouped into pairs (i.e., base and derived word). One pair consisted of the base word and a derived semantically transparent complex word (*n* = 314 pairs; e.g., *comfort–comfortable*), and the other pair was formed by the base word and a derived complex word semantically opaque in relation to its base (*n* = 312 pairs; e.g., *audit–audition*). Phonological transparency was not considered in this selection. In Spanish the selection of words in trios was possible, and therefore 184 base words, 184 semantically transparent, and 184 semantically opaque Spanish words were compiled, irrespective of phonological transparency. We ensured that all derived words were etymologically related to their base word and also that derived words had a recognizable (i.e., familiar and identifiable) suffix, as listed in Quirk, Greenbaum, Leech, and Svartvik (1985) and the Dictionary of the Spanish Real Academy (Real Academia Española, 2001).

Procedure

Age of acquisition (AoA) AoA values for English and Spanish were estimated using a 7-point scale, where 1 meant *having learnt the word before the age of two*; 2, *between the ages of 3 and 4*; 3, *between the ages of 5 and 6*; 4, *between the ages of 7 and 8*; 5, *between the ages of 9 and 10*; 6, *between the ages of 11 and 12*; and 7, which indicated *having learnt the word at 13 years and older* (Carroll & White, 1973).

Imageability Imageability ratings were obtained for all the Spanish and English words in the same manner. The rating instructions employed by Gilhooly and Logie (1980) were

Table 1 Examples of semantically and phonologically related base and derived words in English and Spanish

	Base Word	Transparent	Opaque
English	explain	explaining	explanatory
Spanish	actuar (to perform)	actuación (performance)	
	abierto (open)		abertura (opening)
semantically related			
English	act	actor	
	apart		apartment
Spanish	actuar (to act)	acto (act)	acta (minutes)

used here. Participants were asked to rate how easy or difficult it was, in their opinion, to create an image for each given word. An 8-point scale was used with 1 representing *I don't know the meaning of this word*; 2, *very hard to evoke a mental image*; 3, *hard to evoke a mental image*; 4, *slightly hard to evoke a mental image*; 5, *neither very easy or difficult to imagine*; 6, *slightly easy to evoke a mental image*; 7, *easy to evoke a mental image*; and finally 8, *very easy to evoke a mental image*.

Semantic distance Semantic distance ratings were obtained for all the Spanish and English words in the same manner. Participants were asked to estimate on a 9-point scale how closely related the meanings of pairs of words were with 1 representing *unrelated meanings*; 2, *very unrelated meanings*; 3, *moderately unrelated meanings*; 4, *slightly unrelated meanings*; 5, *neither related nor unrelated*, 6, *slightly related meanings*; 7, *moderately related meanings*; 8, *very related meanings*; and 9, *totally related meanings*. Participants had the option to indicate that the word was unknown to them. Examples of related words and unrelated words were provided, and participants were encouraged to use the entire scale accordingly.

Word frequency The frequency measures presented here were taken from SUBTLEXus (Brysbaert & New, 2009), for the English words, and from SUBTLEX-ESP (Cuetos et al., 2011), for the Spanish words. Both corpuses are based on language samples from subtitles (51 million words in the case of SUBTLEXus and 41.5 million words in the case of SUBTLEX-ESP).

Brysbaert and New (2009) showed that these frequency estimations account for a higher proportion of the variance in naming speeds than do the more traditional frequency values taken from Kučera and Francis (1967) and CELEX (Baayen, Piepenbrock, & van Rijn, 1993), or than other, more contemporary databases, such as HAL (Lund & Burgess, 1996) and Zeno, Ivens, Millard, and Duvvuri (1995). The calculated values per million words, present in both databases, were taken into account for the present norms.

Word length

Numbers of phonemes, letters, and syllables Phonological (number of phonemes and number of syllables) and orthographic (number of letters) measures of word length were computed. Numbers of phoneme counts for each morphologically complex and base word in the present database were calculated following the phonetic characteristics of the English and Spanish languages.

Similarly, numbers of syllables and letters were computed accordingly.

Number of orthographic neighbors The number of orthographic neighbors was defined as the number of words in each language that could be generated by changing one single letter of each target word, while keeping the place of the remaining letters unchanged (Coltheart et al., 1977).

Morphological family size

Morphological family size counts were computed from the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995). The morphological family size for English words was computed as in De Jong et al. (2000), which was based on the number of occurrences of a morpheme constituent of a given simplex word form in its complex derived forms (excluding tokens and inflectional variants). For Spanish, the morphological family size was computed from www.gedlc.ulpgc.es/investigacion/scogeme02/relmorfo.htm, a search database in which the user can look for all of the morphological relations to a given word or can specify a semantic relationship, among other things.

Two measures of morphological family size were computed: (1) the *total* morphological family size, based on the number of complex word types (lemmas, and compounds) that could be derived from a given base word, and (2) the morphological family size of only *derived words that were semantically related to the base form* (De Jong et al., 2000).

Results

The ratings collected were collapsed across lists within each language for AoA, imageability, and semantic relatedness estimations. Descriptive statistics for each variable in each morphological class and language are presented in Table 2.

All of the words and normative values, organized by language and morphological class, are presented in the Online Appendices (A to F). Two correlation matrices, one for English and one for Spanish, are presented in Table 3 A and B, respectively, to ensure that the significance of the correlations reported was meaningful and valid. The data were appropriately transformed to deal with skewed distributions. Thus, the logarithm transformation of word frequency was used.

Some of the correlations in Table 3 are particularly interesting. For example, AoA correlated highly with all the measures of word length in both languages. This supports the developmental hypothesis suggested by Reilly and Kean (2007), in which English base words, free from affixes and therefore shorter, are learned earlier than complex and longer words. Also in both languages, imageability shows lower

Table 2   Descriptive statistics for the words in our database

	Base Word		Phonologically Transparent		Phonologically Opaque		Semantically Transparent		Semantically Opaque	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
A. ENGLISH										
Phonological										
AoA	3.6	1.1	4.3	0.9	4.3	1.0				
Imageability	5.3	1.2	4.9	1.1	4.9	1.0				
Word frequency[*]	3.0	0.6	2.1	0.8	2.4	0.8				
Number of letters	5.9	1.7	8.0	1.8	7.5	2.5				
Number of Phonemes	4.9	1.7	6.8	1.9	6.4	2.4				
Number of syllables	1.9	0.8	2.6	0.8	2.5	1.3				
N	4.2	5.5	2.0	2.7	2.9	5.0				
Semantic										
AoA	3.9	1.2					4.6	0.9	5.4	1.1
Imageability	5.7	1.4					5.3	1.3	4.5	1.7
Word frequency[*]	2.6	0.8					1.8	0.8	1.4	1
Number of letters	5	1.2					7.8	1.6	7.2	1.5
Number of Phonemes	4.3	1.3					6.3	1.7	6.1	1.7
Number of syllables	1.4	0.6					2.6	0.8	2.5	0.7
N	5.4	5.8					1.3	2	1.6	1.9
Semantic relation to base word	–	–					8.7	0.9	4.8	1.8
B. SPANISH										
Phonological										
AoA	3.1	1.2	4.2	1.1	4.9	1.1				
Imageability	5.9	1.2	6.4	1.1	6.0	1.0				
Word frequency[*]	2.9	0.9	1.8	0.9	1.6	0.9				
Number of letters	5.6	1.4	7.6	1.3	7.7	1.6				
Number of Phonemes	5.5	1.4	7.4	1.3	7.6	1.5				
Number of syllables	2.2	0.6	3.3	0.6	3.2	0.7				
N	5.3	5.6	2.7	3.2	2.0	2.5				
Semantic										
AoA	3.8	1.4					4.8	1.1	4.7	1.2
Imageability	6.0	1.2					4.9	1.1	4.8	1.4
Word frequency[*]	2.7	1.0					1.8	0.9	1.8	1.0
Number of letters	6.1	1.5					8.2	1.7	7.8	1.8
Number of Phonemes	6.0	1.5					8.1	1.7	7.7	1.8
Number of syllables	2.5	0.6					3.5	0.7	3.3	0.8
N	4.7	6.3					1.6	2.4	2.0	3.1
Semantic relation to base word	–	–					6.5	1.2	4.1	1.2

[*] Frequency values for English and Spanish are based on estimations per million words.

correlations with the rest of the variables, apart from AoA. However, the correlations are all significant, indicating that there is a tendency for longer and therefore morphologically complex words to be less imageable. Both measures of morphological family size correlated negatively with AoA and word length in both languages, supporting the view that words with much larger morphological family sizes are acquired much earlier and are shorter in length (Henry & Kuperman, 2013). Furthermore, both measures of morphological family size correlated positively, again in both languages, with imageability and word frequency, suggesting that words with larger morphological family sizes are more imageable and occur more frequently. The total morphological family size correlated positively with number of orthographic neighbors for both languages. The correlations support recent findings in which larger morphological family sizes have been shown to

Table 3 Correlation matrix for seven variables with all of the English and Spanish words

	Imag	Freq	No of Lett	No of Phon	No of Syll	No of N	MFS Total	MFS Sem
A. English								
Age of acquisition	−.63**	−.75**	.49**	.47**	.50**	−.46**	−.13**	−.20**
Imageability		.41**	−.32**	−.32**	−.36**	.26**	.10**	.11**
Word frequency			−.36**	−.33**	−.36**	.34**	.14**	.24**
Number of letters				.91**	.84**	−.64**	−.12**	−.12**
Number of phonemes					.84**	−.59**	−.13**	−.00
Number of syllables						−.57**	−.12**	−.02
Number of orthographic neighbours							.09**	−.04
Morphological family size total								.66**
B. Spanish								
Age of acquisition	−.63**	−.66**	.43**	.43**	.41**	−.35**	−.13**	−.26**
Imageability		.33**	−.25**	−.27**	−.23**	.26**	.13**	.29**
Word frequency			−.39**	−.38**	−.41**	.27**	.06*	.11**
Number of letters				.98**	.87**	−.57**	−.13**	−.24**
Number of phonemes					.87**	−.57**	−.14**	−.25**
Number of syllables						−.45**	−.15**	−.23**
Number of orthographic neigbours							.19**	−.20**
Morphological family size total								−.65**

Imag = imageability; Freq = frequency, No of Lett = Number of letters; No of Syll = Number of syllables; No of N = Number of orthographic neighbors; MFS Total = Total morphological family size; MFS Sem = Morphological family size of semantically derived words only.

facilitate word recognition in adults, in students in their second grade, but not in students in their fourth grade (Perdijk et al., 2012). Perdijk et al. suggested that the lack of facilitation with fourth graders corresponds to a specific developmental stage.

One-way analyses of variance and independent t tests were used to compare the values of the words in each morphological category (i.e., base, transparent, and opaque) on AoA, imageability, and word frequency. Different analyses were run for each language and for each set of morphologically related words (i.e., based on phonology or semantics). There were significant differences among the morphological categories for the three variables in both languages. These were analyzed further, when needed, by using post-hoc tests (Tukey's HSD) to compare the categories pairwise on each of the factors (see the summary in Table 4).

Age of acquisition

Base, transparent, and opaque English words phonologically related differed in the AoA ratings given to them [$F(2, 935)=$ 50.88, $MSE=$ 51.37, $p<$ .001]. Post-hoc tests showed that AoA ratings for base words were significantly lower ($M=$ 3.6, $SD=$ 1.1) than those of transparent ($M=$ 4.3, $SD=$ 0.9) and opaque ($M=$ 4.3, $SD=$ 1.0) words. However, transparent and opaque words were not rated as being acquired at significantly different ages. Independent t tests were used to analyze

semantically related English words. Bonferroni correction was applied, adjusting alpha levels to .016 (.05/3). It was found that English base words were rated significantly lower in AoA than English semantically transparent words [$t(624)=$ −12.19, $p<$ .001] and English semantically opaque words [$t(622)=$ −12.81, $p<$ .001]. In addition, English semantically

Table 4 Analysis of variance results for age of acquisition and imageability ratings of words in our database

	English	Spanish
Age of Acquisition Ratings		
Phonological relation	Base < Transparent**	Base < Transparent**
	Base < Opaque**	Base < Opaque**
	Transparent = Opaque	Transparent< Opaque**
Semantic relation	Base < Transparent**	Base < Transparent**
	Base < Opaque**	Base < Opaque**
	Transparent < Opaque**	Transparent= Opaque
Imageability Ratings		
Phonological relation	Base >Transparent**	Base < Transparent**
	Base > Opaque**	Base = Opaque
	Transparent = Opaque	Transparent> Opaque**
Semantic relation	Base >Transparent**	Base > Transparent**
	Base > Opaque**	Base > Opaque**
	Transparent > Opaque**	Transparent = Opaque

> indicates "rated significantly higher than"; < indicates "rated significantly lower than."

transparent words were rated significantly lower than English semantically opaque words [$t(623)= -10.06, p< .001$].

Similar results were observed in the Spanish language. Bonferroni correction was applied, adjusting alpha levels to .016 (.05/3). Thus, independent t tests showed that Spanish base words had significantly lower ratings on AoA than Spanish phonologically transparent words [$t(236)= -5.32, p< .001$] and phonologically opaque words [$t(236)= -15.30, p< .001$]. Unlike in English, Spanish complex and phonologically transparent words were rated significantly lower on AoA than were Spanish complex and phonologically opaque words [$t(236)= 5.25, p< .001$]. Finally, base, transparent, and opaque Spanish semantically related words differed in the AoA ratings given to them [$F(2, 551)= 36.96, MS= 58.12, p< .001$]. Post-hoc tests showed that AoA ratings for Spanish base words were significantly lower ($M= 3.8, SD= 1.4$) than those for transparent ($M= 4.8, SD= 1.1$) and opaque ($M= 4.7, SD= 1.2$) words. However, transparent and opaque words were not rated as being acquired at significantly different ages.

Imageability

English base and complex phonologically related words differed significantly in imageability ratings [$F(2, 935)= 18.55, MSE= 22.55, p< .001$]. Post-hoc tests showed that base words ($M= 5.3, SD= 1.2$) were rated as being significantly more imageable than phonologically transparent ($M= 4.9, SD= 1.1$) and opaque ($M= 4.9, SD= 1.1$) words. However, transparent and opaque words were rated as being equally imageable. Independent t tests showed that semantically related English base words were rated as being more imageable than transparent words [$t(624)= 5.07, p< .001$] and opaque words [$t(622)= 8.07, p< .001$]. In turn, semantically related transparent words were rated as being more imageable than opaque words [$t(623)= 6.87, p< .001$]. Bonferroni correction was applied, adjusting alpha levels to .016 (.05/3).

In Spanish, independent t tests showed that base words had significantly lower ratings for imageability than Spanish phonologically transparent words [$t(236)= -4.09, p< .001$]. Base and complex phonologically opaque words did not differ significantly in their imageability ratings [$t(236)= 0.233, p> .1$], but phonologically transparent words were rated as being more imageable than opaque words [$t(236)= -3.24, p< .001$]. Bonferroni correction was applied, adjusting alpha levels to .016 (.05/3).

Finally, base, transparent, and opaque Spanish semantically related words differed in the imageability ratings given to them [$F(2, 551)= 52.58, MSE= 58.16, p< .001$]. Post-hoc tests showed that imageability ratings for Spanish base words were significantly higher ($M= 6.0, SD= 1.2$) than those of transparent ($M= 4.9, SD= 1.1$) and opaque ($M= 4.8, SD= 1.4$) words. However, transparent and opaque words were not significantly different in their imageability ratings.

Discussion

The aim of the present study was to generate norms for base and morphologically complex words that are phonologically transparent, phonologically opaque, semantically transparent, and semantically opaque, in two languages with slight differences in their morphological structures: English and Spanish. Our view is that the wide range of data provided here, on two of the most commonly spoken languages in the world, could allow fruitful within- and cross-linguistic comparisons.

In addition to the morphological classification of the words presented in these norms, six factors known to affect simple and complex words processing were included. These were AoA, imageability, word frequency, word length, and lexical similarity, as measured by number of orthographic neighbors and morphological family size. The data were explicitly collected for AoA and imageability on all of the words. In addition, semantic similarity was rated for those words classified as semantically related.

The correlations between the variables considered in the present study were highly similar across languages. As has often been found, AoA showed high correlations with imageability and word frequency (e.g., Morrsion & Ellis, 1995). Interestingly, AoA also correlated highly with all of the measures of word length in both languages, supporting the developmental hypothesis (Reilly & Kean, 2007), which proposes that shorter words (i.e., base words) are learned earlier than longer words (i.e., complex words). In addition, imageability showed significant correlations with all of the variables, indicating that there is a tendency for longer and therefore morphologically complex words to be less imageable, as was also suggested by Reilly and Kean. Total morphological family size also showed significant correlations with all other variables considered across the two languages, showing that words with large family sizes tend to be acquired earlier (Henry & Kuperman, 2013), more imageable, more frequent, and shorter in length and to have a larger number of orthographic neighbors. This observation certainly supports the view that there may be semantic involvement (De Jong, 2002; De Jong et al., 2000; Henry & Kuperman, 2013; Steyvers & Tenebaum, 2005) in the manner in which morphological families develop in the mental lexicon. Steyvers and Tenenbaum posited the influence of AoA in a growing semantic network in the mental lexicon. It has been suggested that semantic nodes that are acquired earlier are likely to have greater and more dominant semantic connections than later-acquired semantic nodes. Later-acquired words are disadvantaged because their meanings are built on those of earlier-acquired words, and thus exhibit a processing cost relative to earlier-acquired words (Steyvers & Tenenbaum, 2005). Furthermore, earlier-acquired words also tend to show larger morphological families (Henry & Kuperman, 2013) and larger semantic relatedness (Schreuder & Baayen, 1997). Thus, the

strength and size of the semantic connections for those words with a larger morphological family size, which are more imageable and have been acquired earlier, will be more robust than those of words with less morphological and semantic connectivity and that have been acquired much later (De Jong, 2002; De Jong et al., 2000; Henry & Kuperman, 2013; Steyvers & Tenenbaum, 2005).

The statistical comparisons between the mean AoA ratings given to base and morphologically complex words, both those phonologically and semantically related, yielded the same pattern of results across languages (see Table 4 for a summary of the results). Base words were consistently rated as being acquired significantly earlier than morphologically complex words. This suggests that the reported influence of the AoA of the base word when identifying the inflected word forms (Kuperman et al., 2012) is more likely to be due a morphological decomposition process present in word recognition than to the possibility that the base and inflected word forms share the same AoA values.

Furthermore, other aspects of the present study require careful consideration in future cross-linguistic research. Our colleague and reviewer Cristina Burani pointed out that the different morphological richnesses of the two languages under study might have an impact on the computed surface frequencies presented in these norms. Thus, in the case of English, a language morphologically simpler than Spanish, surface frequencies mirror cumulative frequencies very closely (e.g., adjectives do not have inflected forms; the surface frequency of base nouns only excludes the plural form; surface frequencies of verb bases only exclude the frequencies of their three inflected forms -s, -ing, and -ed; etc.), In contrast, surface frequencies of most Spanish words exclude a good number of inflected forms (e.g., plural and gender for adjectives, a high number of tenses [past, past participle, present, future, conditional, etc.], and declinations depending on the corresponding pronouns for verbs). This implies that the surface frequencies of the English language are likely to be higher than those of the Spanish language. We tested this by randomly selecting 20 English and 20 Spanish verbs from the phonologically related words and running a t test. The results showed a statistically significant difference [$t(38) = -3.59$, $p < .001$], with English base verbs having on average a higher frequency than Spanish base verbs. This is a very interesting result that deserves future exploration. The fundamental morphological differences between English and Spanish affect the surface frequencies of base words in English and Spanish base words, and this need to be taken into account in future cross-language studies.

In summary, as we previously discussed, very few (if any) studies of morphological processing have considered AoA; therefore, the present study is an attempt to fill a crucial gap in morphological processing research. We provide a comprehensive compilation of ratings for AoA, imageability, and semantic transparency. To add to this compilation, supplementary normative data were derived from databases for frequency, word length, number of orthographic neighbors, and morphological family size. In addition, in order to facilitate the selection of stimuli in future studies, the nature of the relationships between the simple and complex words has been divided between phonological or semantic. Furthermore, comparisons between the ratings provided for the base and complex word forms on imageability in the present study largely support Reilly and Kean's (2007) proposal that complex words are less concrete, and therefore less imageable, than simple or base words. Intriguingly, and also in need of further investigation, the Spanish base and complex words that were phonologically related did not follow such a pattern, with phonologically transparent words showing the higher imageability ratings.

Indeed, the ratings provided in the present study for two extensively spoken languages offer a unique methodological enterprise and contribution to the future study of morphological processing.

References

Alija, M., & Cuetos, F. (2006). Effects of the semantic variables in visual word recognition. *Psicothema, 18,* 485–491.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 802–814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 234–254. doi:10.1037/0278-7393.18.2.234

Assink, E. M. H., van Well, S., & Knuijt, P. P. N. A. (2003). Age-of-acquisition effects in native speakers and second-language learners. *Memory & Cognition, 31,* 1218–1228.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language, 36,* 94–117. doi:10.1006/jmla.1997.2509

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55,* 290–313. doi:10.1016/j.jml.2006.03.008

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (Release 2, CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316. doi:10.1037/0096-3445.133.2.283

Bertram, R., Baayen, R. H., & Schreuder, R. (2002). Effects for family size for complex words. *Journal of Memory and Language, 42,* 320–405.

Bonin, P., Chalard, M., Méot, A., & Barry, C. (2006). Are age-of-acquisition effects on object naming due simply to differences in object recognition? Comments on Levelt (2002). *Memory & Cognition, 34,* 1172–1182. doi:10.3758/BF03193263

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,* 545–559. doi:10.1080/17470218.2010.503374

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

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica, 104,* 215–226. doi:10.1016/S0001-6918(00)00021-4

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes, 2,* 217–227.

Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language, 18,* 342–352.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28,* 297–332.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology, 25,* 85–95. doi:10.1080/14640747308400325

Catling, J. C., Dent, K., & Williamson, S. (2008). Age of acquisition, not word frequency affects object recognition: Evidence from the effects of visual degradation. *Acta Psychologica, 129,* 130–137.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11,* 63–65.

Christianson, K., Johnson, R. L., & Rayner, K. (2005). Letter transpositions within and across morphemes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1327–1339. doi:10.1037/0278-7393.31.6.1327

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3,* 149–170.

Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

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. doi:10.1037/0033-295X.108.1.204

Connine, C. M., Mullenix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1084–1096. doi:10.1037/0278-7393.16.6.1084

Criss, A. H., Aue, W., & Smith, L. (2011). The effects of word frequency and context variability in cued recall. *Journal of Memory and Language, 64,* 119–132.

Cuetos, F., Barbón, A., Urrutia, M., & Domínguez, A. (2009). Determining the time course of lexical frequency and age of acquisition using ERP. *Clinical Neurophysiology, 120,* 285–294. doi:10.1016/j.clinph.2008.11.003

Cuetos, F., González-Nosti, M., Barbón, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica, 32,* 133–143.

Cuetos, F., Samartino, T., & Ellis, A.W. (2012). Age acquisition norms from elderly Spanish people: Characteristics and the prediction of word recognition performance in Alzheimer's disease. *Psicológica, 33,* 59–76. Retrieved from www.uv.es/psicologica/articulos1.12/4CUETOS.pdf

Cutler, A., Hawkins, J. A., & Gillian, G. (1985). The suffixing preference: A processing explanation. *Linguistics, 23,* 723–750.

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

De Jong, N. H. (2002). *Morphological families in the mental lexicon (MPI Series in Psycholinguistics)*. Nijmegen, The Netherlands: Max Planck Institute of Psycholinguistics.

De Jong, N. H., Shreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes, 15,* 329–365. doi:10.1080/01690960050119625

Devlin, J. T., Jamison, H. L., Matthews, P. M., & Gonnerman, L. M. (2004). Morphology and the internal structure of words. *Proceedings of the National Academy of Sciences, 101,* 14984–14988.

Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology, 25,* 108–127.

Dijkstra, T., del Prado, M., Martin, F., Schulpen, B., Schreuder, R., & Baayen, R. H. (2005). A roommate in cream: Morphological family size effects on interlingual homograph recognition. *Language and Cognitive Processes, 20,* 7–41.

Dimitropoulou, M., Duñabeitia, J. A., Avilés, A., Corral, J., & Carreiras, M. (2010). Subtitle-based word frequencies as the best estimate of reading behavior: The case of Greek. *Frontiers in Psychology, 1,* 218. doi:10.3389/fpsyg.2010.00218

Domínguez, A., Seguí, J., & Cuetos, F. (2002). The time-course of inflectional morphological priming. *Linguistics, 40,* 235–259.

Duñabeitia, J. A., Laka, I., Perea, M., & Carreiras, M. (2009). Is *Milkman* a superhero like *Batman*? Constituent morphological priming in compound words. *European Journal of Cognitive Psychology, 21,* 615–640. doi:10.1080/09541440802079835

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2007). Do transposed-letter similarity effects occur at a morpheme level? Evidence for ortho-morphological decomposition. *Cognition, 105,* 691–703. doi:10.1016/j.cognition.2006.12.001

Ellis, A. W., Burani, C., Izura, C., Bromiley, A., & Venneri, A. (2006). Traces of vocabulary acquisition in the brain: Evidence from covert object naming. *NeuroImage, 33,* 958–968.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge: MIT Press.

Española, R. A. (2001). *Diccionario de la lengua española [Dictionary of the Spanish Language]* (22nd ed.). Madrid: Espasa Calpe.

Ferrand, L., Brysbaert, M., Keuleers, E., New, B., Bonin, P., Méot, A., & Pallier, C. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in Psychology, 2,* 306. doi:10.3389/fpsyg.2011.0036

Fowler, H. W., & Fowler, F. G. (1982). *The concise Oxford dictionary of current English* (7th ed.). Oxford: Oxford University Press, Clarendon Press.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 361–379. doi:10.1037/0096-1523.2.3.361

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition, 28,* 1277–1288. doi:10.3758/BF03211828

Gilhooly, K. J., & Gilhooly, M. L. M. (1980). The validity of age of acquisition ratings. *British Journal of Psychology, 71,* 105–110. doi:10.1111/j.2044-8295.1980.tb02736.x

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12,* 395–427. doi:10.3758/BF03201693

Gonnerman, L. M., & Plaut, D. C. (2000). *Semantic and morphological effects in masked priming. Poster presented at the annual meeting of the Cognitive Neuroscience Society.* San Francisco: CA.

Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General, 136,* 323–345. doi:10.1037/0096-3445.136.2.323

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111,* 662–720.

Henry, R., & Kuperman, V. (2013). Semantic growth of morphological families in English. *Psihologija, 46,* 479–495.

Hirsh, K. W., & Ellis, A. W. (1994). Age of acquisition and aphasia: A case study. *Cognitive Neuropsychology, 11,* 435–458.

Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition, 6,* 117–128.

Holmes, S., Fitch, F. J., & Ellis, A. W. (2006). Age of acquisition affects object recognition and naming in patients with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 28,* 1010–1022.

Hudson, P. T. W., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory and Language, 24,* 46–58.

Izura, C., & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica, 23,* 245–281.

Izura, C., & Ellis, A. W. (2004). Age of acquisition effects in translation production tasks. *Journal of Memory and Language, 50,* 165–181.

Izura, C., Hernández-Muñoz, N., & Ellis, A. W. (2005). Category norms for 500 Spanish words in five semantic categories. *Behavior Research Methods, 37,* 385–397. doi:10.3758/BF03192708

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,* 32–58. doi:10.1016/j.jml.2010.09.002

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

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrlated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1312–1318. doi:10.1037/0278-7393.29.6.1312

Juhasz, B. J., & Rayner, K. (2006). The role of age-of-acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition, 13,* 846–863.

Kennet, J., McGuire, L., Willis, S. L., & Shaie, K. W. (2000). Memorability functions in verbal memory: A longitudinal approach. *Experimental Aging Research, 26,* 121–137.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American English.* Providence, RI: Brown University Press.

Kuperman, V., Bertram, R., & Baayen, R. H. (2010). Processing trade-offs in the reading of Dutch derived words. *Journal or Memory and Language, 62,* 83–97.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44,* 978–990. doi:10.3758/s13428-012-0210-4

Lavidor, M., & Ellis, A. W. (2002). Orthographic neighborhood effects in the right but not in the left cerebral hemisphere. *Brain and Language, 80,* 63–76.

Lavric, A., Rastle, K., & Clapp, A. (2011). What do fully visible primes and brain potentials reveal about morphological decomposition? *Psychophysiology, 48,* 676–686.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge: MIT Press.

Liu, Y., Hao, M., Shu, H., Tan, L. H., & Weekes, B. S. (2008). Age-of-acquisition effects on oral reading in Chinese. *Psychonomic Bulletin & Review, 15,* 344–350. doi:10.3758/PBR.15.2.344

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

Macizo, P., & Bajo, M. T. (2006). Reading for repetition and reading for translation: Do they involve the same processes? *Cognition, 99,* 1–34.

MacLeod, C. M., & Kampe, K. E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 132–142. doi:10.1037/0278-7393.22.1.132

Marslen-Wilson, W. D., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form and meaning. *Language and Cognitive Processes, 23,* 394–421.

Marslen-Wilson, W. D., & Tyler, L. K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Sciences, 2,* 428–435.

Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101,* 3–33. doi:10.1037/0033-295X.101.1.3

Mathey, S. (2001). L'influence du voisinage orthographique lors de la reconnaissance des mots écrits. *Revue Canadienne de Psychologie Expérimentale, 55,* 1–23.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences, 6,* 464–465.

Menenti, L., & Burani, C. (2007). What causes the effect of age of acquisition in lexical processing? *Quarterly Journal of Experimental Psychology, 60,* 652–660.

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling–sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 183–206. doi:10.1037/0278-7393.28.1.183

Monaghan, P., & Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language, 63,* 506–525. doi:10.1016/j.jml.2010.08.003

Morris, J., Frank, T., Grainger, J., & Holcomb, P. J. (2007). Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology, 44,* 506–521.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 116–133. doi:10.1037/0278-7393.21.1.116

Mulder, K., Dijkstra, T., Schreuder, R., & Baayen, H. R. (2014). Effects of primary and secondary morphological family size in monolingual and bilingual word processing. *Journal of Memory and Language, 72,* 59–84.

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

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,* 45–52. doi:10.3758/BF03193811

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17,* 273–281.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart & Winston.

Perdijk, K., Schreuder, R., Baayen, H., & Verhoeven, L. (2012). Effects of morphological family size for young readers. *British Journal of Developmental Psychology, 30,* 432–445.

Perea, M., & Rosa, E. (2000). Repetition and form priming interact with neighborhood density at a short stimulus-onset asynchrony. *Psychonomic Bulletin & Review, 7,* 668–677.

Pérez, M. A. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *Quarterly Journal of Experimental Psychology, 60,* 32–42. doi:10.1080/17470210600577423

Piera, C. (1995). On compounding in English and Spanish. In H. Campos & P. Kempchinsky (Eds.), *Evolution and revolution in linguistic theory* (pp. 302–315). Washington, DC: Gerogetown University Press.

Pinker, S. (1991). Rules of language. *Science, 253,* 530–535. doi:10.1126/science.1857983

Pinker, S., & Ullman, M. (2002). The past and future of the past tense. *Trends in Cognitive Sciences, 6,* 456–463.

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15,* 445–485.

Prado, E., & Ullman, M. T. (2009). Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology: Language, Memory, and Cognition, 35,* 849–866.

Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes, 8,* 1–56. doi:10.1080/01690969308406948

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

Raman, I. (2006). On the age of acquisition effects in word naming and orthographic transparency: Mapping specific or universal? *Visual Cognition, 13,* 1044–1053.

Ramirez, G., Chen, X., Geva, E., & Yang, L. (2011). Morphological awareness and word reading in English language learners: Evidence from Spanish- and Chinese-speaking children. *Applied Psycholinguistics, 32,* 601–618.

Rastle, K., & Davis, M. H. (2003). Reading morphologically complex words: Some thoughts from masked priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: State of the art* (pp. 279–305). New York: Psychology Press.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes, 15,* 507–537. doi:10.1080/01690960050119689

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review, 11,* 1090–1098. doi:10.3758/BF03196742

Raveh, M., & Rueckl, J. G. (2000). Equivalent effects of inflected and derived primes: Long-term morphological priming in fragment completion and lexical decision. *Journal of Language and Memory, 42,* 103–119.

Reilly, J., & Kean, J. (2007). Morphophonological distinctiveness of high- and low-imageability nouns: A corpus analysis. *Cognitive Science, 31,* 1–12.

Richards, R. M., & Ellis, A. W. (2009). Identity, gender, and the role of age of acquisition in face processing. *Psicológica, 30,* 155–179.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale: Erlbaum.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language, 37,* 118–139.

Sears, C. R., Hino, Y., & Lupker, S. J. (1999). Orthographic neighbourhood effects in parallel distributed processing models. *Canadian Journal of Experimental Psychology, 53,* 220–230.

Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences, 4,* 353–361.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the ítem level. *Psychological Science, 8,* 411–416. doi:10.1111/j.1467-9280.1997.tb00453.x

Stanners, R. F., Neiser, J. J., Hernon, W. P., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior, 18,* 399–412. doi:10.1016/S0022-5371(79)90219-6

Stewart, N., & Ellis, A. W. (2008). Order of acquisition in learning perceptual categories: A laboratory analogue of the age-of-acquisition effect? *Psychonomic Bulletin & Review, 15,* 70–74. doi:10.3758/PBR.15.1.70

Steyvers, M., & Tenenbaum, J. B. (2005). The large scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29,* 41–78. doi:10.1207/s15516709cog2901_3

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1140–1154. doi:10.1037/0278-7393.21.5.1140

Taft, M. (1979). Lexical access via an orthographic code: The basic orthographic syllabic structure (BOSS). *Journal of Verbal Learning and Verbal Behavior, 18,* 21–39. doi:10.1016/S0022-5371(79)90544-9

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9,* 271–294.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology, 57A,* 745–765.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647. doi:10.1016/S0022-5371(75)80051-X

Tree, J. J., Perfect, T. J., Hirsh, K. W., & Copstick, S. (2001). Deep dysphasic performance in non-fluent progressive aphasia: A case study. *Neurocase, 7,* 473–487.

Tyler, L. K., de Mornay Davies, P., Anokhina, R., Longworth, C., Randall, B., & Marslen-Wilson, W. D. (2002). Dissociations in processing past tense morphology: Neuropathology and behavioral studies. *Journal of Cognitive Neuroscience, 14,* 79–94.

Van Hell, J. G., & Dijkstra, A. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9,* 780–789. doi:10.3758/BF03196335

Van Heuven, W. J. B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39,* 458–483.

Vannest, J., Newport, E. L., Newman, A. J., & Bavelier, D. (2011). Interplay between morphology and frequency in lexical access: the case of the base frequency effect. *Brain Research, 1373,* 144–159.

Weber, G. (1997). The world's 10 most influential languages. *Language Today, 2,* 12–18.

Weekes, B. S., Chan, A. H., & Tan, L. H. (2008). Effects of age of acquisition on brain activation during Chinese character recognition. *Neuropsychologia, 46,* 2086–2090. doi:10.1016/j.neuropsychologia.2008.01.020

West, M. (1953). *A general service list of English words*. London: Longman, Green, & Co.

Whaley, C. (1978). Word–nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17,* 143–154.

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, 1842–1859. doi:10.1037/a0033090

Wilson, M. A., Ellis, A. W., & Burani, C. (2012). Age-of-acquisition affects word naming in Italian only when stress is irregular. *Acta Psychologica, 139,* 417–424. doi:10.1016/j.actpsy.2011.12.012

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441–517. doi:10.1006/jmla.2002.2864

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster: Touchstone Applied Science Associates.