Psycholinguistic norms for a set of 506 French compound words

Patrick Bonin[1] · Betty Laroche[1] · Alain Méot[2]

Abstract

Compounds are morphologically complex words made of different linguistic parts. They are very prevalent in a number of languages such as French. Different psycholinguistic characteristics of compounds have been used in certain studies to investigate the mechanisms involved in compound processing (see Table 7). We provide psycholinguistic norms for a set of 506 French compound words. The words were normed on seven characteristics: lexeme meaning dominance, semantic transparency, sensory experience, conceptual familiarity, imageability, age of acquisition (AoA) and subjective frequency. Reliability measures were computed for the collected norms. Descriptive statistical analyses, and correlational and multiple regression analyses were performed. We also report some comparisons made between our normative data and certain norms obtained in other similar studies. The entire set of norms, which will be very useful to researchers investigating the processing of compounds, is available as Supplemental Material.

Compounds are special—morphologically complex—words (Arcara et al., 2014) composed of different linguistic parts (referred to as "lexemes" or "morphological components" or "constituents"). These parts are used to communicate one specific meaning (e.g., *hot dog* meaning a kind of fast food) while the different constituents also individually relate to that meaning to some degree (*hot* and *dog* are differentially related to the meaning of *hot dog*). Compounds differ from derived and inflected words in that they are formed from two free lexemes (Juhasz et al., 2003). An interesting definition of compound words is that provided by Desrochers et al. (2010): "*A word is said to be derived by compounding when two lexemes (rarely more) are joined (e.g., flycatcher). Thus, compound words arise from the concatenation of distinct lexemes, which have their own frequency of occurrence in print and/or speech*" (p. 109).

A very productive way to construct novel words is word compounding (Fiorentino et al., 2014) which is very prevalent in a number of languages (Libben, 2014). To give a few examples: in English, such words are fairly common (about 4800 (token) occurrences of nominal compounds as reported by Janssen et al., 2008), but the German language also makes extensive use of compounds (Lüttmann et al., 2011). The same is true for Basque, which contains a large number of compounds and where compounding is a very often used morphological mechanism (Duñabeitia, Laka, et al., 2009a). In non-alphabetic languages, compounding is also very often employed—for example in Chinese[1] (Cui et al. (2017). There are also compounds in languages that are less known by Westerners, such as the Bangla language (Dasgupta et al., 2016). Finally, as far as French is concerned—the language for which norms of compounds were collected in the present study—we are not aware of any published studies that have provided detailed statistics about the number of compounds. As in other languages, compounds are also expressions used in French to communicate meanings. In sum, compounding can be thought of as a universal way to form morphologically complex words across languages (Dressler, 2006).

✉ Patrick Bonin
  Patrick.Bonin@u-bourgogne.fr

[1] LEAD-CNRS (UMR5022), Univ. Bourgogne Franche-Comté, Pôle AAFE - Esplanade Erasme, BP 26513, 21065 Dijon Cedex, France

[2] LAPSCO-CNRS (UMR6024), Université Clermont Auvergne, Clermont-Ferrand, France

---

[1] According to Lim et al. (2020), there are several varieties or dialects in Chinese: Mandarin Chinese is spoken in mainland China, Taiwan and Singapore whereas Cantonese Chinese is spoken in Hong Kong and Macau.

At the surface level, compounds can take different forms, and the variations seen in forms mostly depend upon the language in question. In German, compounds are consistently written as a single string/single word (Günther et al., 2020; Zwitserlood et al., 2002). In English, compounds are made of free lexemes, and typically a combination of two or more constituent words (Dressler, 2006), which results in single lexicalized expressions (Juhasz et al., 2003)—compare for example *tree house* in English with *Baumhaus* in German. Most English compounds are written as unified expressions whereas others are written with a space or hyphenated (Inhoff et al., 2008). In a similar way, in Korean, compounds are built by combining two or more stem morphemes (Ko et al., 2011), with the noun-noun combination being a productive type of compounding (Sohn, 1999 cited in Ko et al., 2011). Compound words in Chinese are often formed by the combination of two or more characters, as in 雪人 [snowman] (Tse et al., 2017). As mentioned by Tse et al. (2017), about 73.6% of modern Chinese words are two-character compound words. In French, compounds can take different forms. For instance, there are noun-noun compounds (e.g., *radio-réveil* [alarm clock]), adjective-noun compounds (e.g., *chauve-souris* [bat]), verb-noun compounds (e.g., *tire-bouchon* [corkscrew]). There are also "noun-preposition-noun" compounds (e.g., *chef de gare* meaning station master, *sac à main* meaning a handbag) which, Nicoladis and Krott (2007) point out, are more frequent than the noun-noun compound type, with "de" and "à" being the most frequent prepositions found in noun-preposition-noun compounds (e.g., *pommes de terre* meaning potatoes, see Nicoladis, 2001).

Different languages also vary as a function of whether their compounds are left- or right-headed. In English, German or Dutch, compound words are right-headed (Koester & Schiller, 2008), that is to say that the second constituent (rightmost) determines the semantic category and the morphosyntactic features of the whole compound—it serves as the head (Juhasz, 2018). The first constituent serves as a modifier which further specifies the meaning as, for instance, in *airport*, which is a kind of port dedicated to travel by air (Günther et al., 2020). In Bulgarian, compounds are always right-headed (Jarema et al., 1999). In both Italian and French, it is possible to have both left- and right-headed compounds (Semenza & Luzzatti, 2014), even though in French, noun–noun compounds are generally left-headed (Jarema et al., 1999). In French noun–adjective compounds, the adjective can be located at the initial or at the final position; thus there are both left-headed and right-headed noun–adjective compounds.

Different views on the processing of compounds

A key issue in psycholinguistics is to determine the degree to which the different parts are involved when accessing the meaning of compounds. For example, the French compound *chauve-souris*—meaning bat—comprises a left-part adjective *chauve* meaning bald and a right-part noun meaning mouse. When reading *chauve-souris*, do readers first access the meaning of *chauve* and then the meaning of *souris*, and then, finally, the meaning of the compound? Or are the meanings of the two lexemes simply bypassed?[2]

The processing of compounds has given rise to a large number of studies. Most of them have been conducted in the field of comprehension, and less so in word production (Gagné & Spalding, 2016a, 2016b; Zwitserlood et al., 2002). Any word reading (or any word production) model must account for the representations and the processes involved in the processing of all types of words, i.e., not only monomorphemic words but also polymorphemic words as well as compounds (e.g., *hot dog*). The findings on compounds are far from being consistent, but they have helped researchers build and constrain different views on compound processing in production or in word reading tasks. Indeed, different views have been put forward to account for how compounds are processed (for a brief overview, Fiorentino & Poeppel, 2007; Kuperman, 2013). Briefly, one view holds that compound words are processed based on their constituent parts—full-decomposition (or full-parsing) models (e.g., Taft & Forster, 1976). In contrast, according to another radically opposite view—full-listing models (e.g., Butterworth, 1983; Stockall & Marantz, 2006)—compounds are processed based on their "full form", that is to say holistically. Mixed views have also been proposed (e.g., Bertram & Hyönä, 2003; Kuperman et al., 2009; Schreuder & Baayen, 1995) and assume that both types of processing may be involved depending on the characteristics of the compound. For instance, in the "late decomposition model" (e.g., Giraudo & Grainger, 2000), processing starts with whole-word representations and access to the constituents occurs afterward, for instance when the relation between the whole word and the compound's constituents is semantically transparent.

---

[2] This issue is analogous to the processing of idiomatic expressions (Bonin, Méot, Boucheix, & Bugaiska, 2018). Here also, different views lead to different predictions concerning the role played by the literal meaning when accessing the figurative meaning of idioms.

Influence of different psycholinguistic variables on compound processing

One way to disentangle different views on compound processing has been to examine the (independent or combined) influence of several characteristics of compounds pertaining to either their constituents (*hot* and *dog* in *hot dog*) or their full form (*hot dog*). As a result, having norms on compounds is very important in order to address issues pertaining to compounds in psycholinguistics, such as how compounds are represented in the brain, as well as details of their processing in both comprehension and production tasks. In the present paper, we aim to provide psycholinguistic norms based on a set of 506 French compound words. To our knowledge, psycholinguistic norms are not available in French for this type of word. We now describe different psycholinguistic characteristics that have been used in certain studies to investigate compound processing and the rationale behind that use.

When describing the different psycholinguistic norms, we will refer to certain studies that have used each variable in question either in order to control it methodologically or statistically, or to investigate its potential effect in language processing. For the sake of concision, and because the aim of our study was to provide norms on compounds in French, we decided to indicate in a table (Table 7 in the Appendix) certain studies which made use of each of the described variables and not to describe exactly how the variable was explored (e.g., experimental tasks employed, direction of the effects observed). Likewise, readers can easily use this table to refer to the specific studies in order to know what the findings are for the variable in question and in what experimental setting they were obtained. We should make it clear that in doing so we did not intend to provide a full list of the studies in this vast field of research. The studies listed in Table 7 should therefore be taken as illustrations, even though it must be stressed that we were careful to include most recent studies together with older ones. The evidence broadly tends to suggest that there are reliable effects of the constituents' characteristics on the processing of compounds. These effects have been observed in a large number of lexical processing tasks (e.g., lexical decision, word reading, spoken and written picture naming), often using different experimental techniques (e.g., implicit priming paradigm, picture-word interference paradigm, EEG and fMRI techniques). Thus, in the course of compound comprehension or production, the compound's constituents are activated to some degree and they influence the processes underpinning these language abilities.

Objective word frequency is a well-known variable in psycholinguistics and one that has proven to be very influential in lexical processing (see Brysbaert et al., 2018 for a recent review). It corresponds to the number of times a word is found in written (or spoken) corpora. It is possible to evaluate the overall frequency of compounds (e.g., the frequency of the word *hot dog*), but also the frequency of their constituents (e.g., the frequency of *hot* and *dog*). For example, in French, the compound word *chauve-souris* has an overall frequency (subtitle frequency) of 5.43 per million, whereas *chauve* and *souris* have frequency values of 5.25 and 21.94 per million, respectively. The rationale has been that, whenever the frequencies of the constituents have an impact on the processing of compounds, this is then evidence for the hypothesis that they are decomposed in some way. One research strategy has been to manipulate the frequency of the constituents of compounds. This strategy was initially undertaken by Taft and Forster (1976) and it has been employed very often since (Table 7).

Length, like lexical frequency, is an objective variable. It has often been examined in visual word recognition but its effects have proven to be non-intuitive (e.g., Ferrand et al., 2011; New et al., 2006). The influence of length has been examined in connection with compounds on the basis of eye movements in reading and it has been found that longer compounds yielded longer gaze durations than shorter compounds (e.g., Juhasz, 2008). Length has been found to interact with frequency in various ways such as, for example, in Bertram and Hyönä's (2003) study, in which constituent frequency effects were found for long but not for short compound words.

Subjective frequency and age of acquisition (AoA) are also lexical variables that have been investigated in compound processing (even though certain researchers think that AoA also indexes semantic aspects, see for instance Ghyselinck et al., 2004). Subjective frequency is a measure of how often a word is encountered in everyday life (read or heard). To assess subjective frequency for words, Likert-like scales are used, for example a 7-point scale where the response 1 is given for words that are perceived as being rare and the response 7 for words that are perceived as being commonly used (Desrochers et al., 2010; Desrochers & Thompson, 2009). Subjective and objective frequency scores are correlated (e.g., Bonin, Méot, et al., 2003a; Desrochers et al., 2010; Gonthier et al., 2009), thus making it possible to complement objective frequency measures (Desrochers & Thompson, 2009), especially for objectively low-frequency words that occur in very specific contexts (e.g., *syringe* is a low-frequency word according to objective frequency counts but is subjectively frequent for people working in hospitals). They can also be used as a proxy when objective frequency measures are lacking for some words. Interestingly, Balota et al. (2001) have shown that subjective frequency ratings are a slightly better predictor of lexical decision times than

objective word frequency measures. An influence of familiarity/subjective frequency on compound processing has been found in word reading (Table 7), but this variable has been less-well investigated than objective lexical frequency.

AoA corresponds to the age at which a word was acquired. AoA of words is often measured by asking adults to rate the age at which they think they learned different words. The AoA effect—words acquired early in life are processed faster and more accurately than words acquired later in life—is a strong empirical finding that surfaces in a large number of lexical tasks (see for reviews: Juhasz, 2005; Johnston & Barry, 2006). This phenomenon is still debated at both the methodological, i.e., how best to evaluate the AoA of words, and theoretical levels, i.e., what are the mechanisms underpinning AoA effects. Another related question is what is the locus of AoA effects (e.g., Bonin et al., 2004; Brysbaert, 2017; Cortese & Khanna, 2007; Zevin & Seidenberg, 2002, 2004; see Calting & Elsherif, 2020 for a brief overview). Up to now, as can be seen from Table 7, the impact of AoA has not often been investigated in compound processing.

Semantic transparency corresponds to the relationship between the meaning of the full form of the compound (e.g., *hot dog* is a piece of food) and the meaning of its constituents (e.g., *hot* refers to temperature and *dog* refers to an animal that barks). As claimed by Günther et al. (2020), research on compounds has made widespread use of the notion of semantic transparency. Compounds vary as a function of their semantic transparency. Fully transparent compounds comprise two lexemes that both contribute to the overall meaning (e.g., *birdhouse*). In contrast, in fully opaque compounds (e.g., *shindig*), neither constituent contributes to the meaning, whereas in partially opaque words (e.g., *hot dog*), only one lexeme is linked to the meaning of the compounds. Thus, there can be varying degrees of semantic transparency in compounds, making it possible to assess the extent to which the meaning of compounds is predictable from the meaning of the constituents. As listed in Table 7, the influence of semantic transparency has been explored in morphological processing in different languages and in several experimental tasks related to either language comprehension or production. Norms of semantic transparency have been collected in English (e.g., Gagné et al., 2019; Kim et al., 2019) and, very recently, semantic transparency measures have been provided for a large set of German compounds (Günther et al., 2020). Another psycholinguistic characteristic that is specific to compounds and for which it is worth collecting norms is *lexeme (or lexical) dominance*. As stated earlier, languages vary as a function of whether their compounds are left- or right-headed. The meaning can thus be carried either by the left or by the right lexeme or by both lexemes. Lexeme dominance can be assessed by asking participants to use Likert scales to evaluate

the degree to which each lexeme contributes to the whole meaning of the compounds. The influence of lexeme dominance in compound processing has been examined in several languages as reported in Table 7.

Apart from the above characteristics, there are other variables that have been found to influence the processing of compounds. Conceptual familiarity, imageability and sensory experience ratings (SER) are all semantic variables which presumably index different semantic dimensions (Kuperman, 2013). Conceptual familiarity measures the degree of physical or mental contact with an object (e.g., a *rake* is very familiar for a gardener). This variable has often been used to investigate the determinants of object naming speed (Perret & Bonin, 2019). However, this variable has not been taken into account very often when investigating visual word recognition (but see Chedid et al., 2019, for a recent study showing a reliable influence of this factor on lexical decision times). We are not aware of any study on compounds which has examined the influence of this variable. Imageability is a variable thought to index the richness of words (Yap & Pexman, 2016) and it corresponds to the ease with which a mental image can be formed in response to words (Bonin et al., 2011). Imageability is a reliable predictor in a large number of word recognition tasks, such as lexical decision, word reading or conditional word naming (e.g., Cortese et al., 2018; Strain & Herdman, 1999), but it has not been investigated to any great extent in compound processing (see Table 7 for a study examining the influence of imageability in lexical decision times in response to compounds).

SERs are thought to reflect the extent to which a word evokes a sensory and/or perceptual experience (Juhasz et al., 2011; Juhasz & Yap, 2013). To provide SER for words, participants use Likert scales to assess the extent to which the words evoke an actual sensation (e.g. sight, touch, smell). Bonin et al. (2015) found that SER reliably predicted response times for lexical decision on individual words, but not in word naming or progressive demasking. SER effects have also been reported for compound processing (Table 7).

Given that no norms on compounds are available in French, the aim of our study was to provide norms for a set of 510 compounds, and these should prove to be very helpful to researchers who wish to investigate the processing of such words.

Collection of the norms

Method

Participants A total of 216 students from bachelor's degree to master's degree at the University of Bourgogne Franche-Comté (mean age: 20.41 years; 44 male) completed the

different questionnaires. All were French native speakers. Thirty-one participants completed each questionnaire, except in the case of the semantic dominance questionnaire, for which there were only thirty participants. This study involving human participants was reviewed and approved by the Statutory Ethics Committee of the University Clermont Auvergne.

Material

The number of French words that have been rated for certain psycholinguistic norms is much smaller than is the case for other languages such as English, even though there are now a growing number of studies that provide psycholinguistic norms in French (see http://www.lexique.org/?page_id=378). We started by checking published French normative studies for their inclusion of compounds. It turned out that compounds were rarely included in the normed items. For instance, there are 45 compounds corresponding to object names in the normative study by Alario and Ferrand (1999) for 400 line drawings taken from Cycowicz et al. (1997), including the 260 Snodgrass and Vanderwart (1980) line drawings. In the Bonin, Peereman, et al. (2003b) normative study of 299 line drawings, only five compounds corresponding to object names can be found. In the Ferrand et al. (2008) study, monosyllabic words ($N = 1{,}493$) were rated on AoA and subjective frequency, and thus, no compounds were included. Finally, in the MEGALEX study (Ferrand et al., 2018), no compounds were included.

Juhasz et al. (2015) selected their stimuli from a larger set of English words (Juhasz & Berkowitz, 2011) and had a list of 629 compound words. Given that compound words are different in English and French, it was not possible to translate the list of compounds provided by Juhasz et al. (2015) in a way similar to how pictures normed in one language can be used directly to collect norms in another. We therefore had to create our own list of compounds. Our objective was to gather a list of compounds that university students (who are often used as participants in psycholinguistic studies) might know, and which could therefore be used in word reading or language production experiments involving undergraduate students. Also, in the same way as Juhasz et al. (2015), we were careful to select compound words that would vary in familiarity. Because, to our knowledge, there is no specific database of compounds in French, we gathered an extensive list of French compounds (both with and without prepositions) by searching through several online dictionaries available on the internet and in the Lexique.org database (New et al., 2004). More precisely, about 400 words from the Lexique database were selected, while taking into account that each compound had to be familiar to undergraduates in order to be included in the list. In the end, 397 compounds were selected. Using the same criterion, the remaining compounds ($N = 113$) were chosen by consulting French dic-

tionaries. Likewise, we selected 510 French compound words of different lengths. Length is defined as the number of characters (spaces, hyphens and apostrophes included).

Procedure Seven different types of norms were collected using the questionnaires: lexeme meaning dominance, semantic transparency, sensory experience rating (SER), conceptual familiarity, imageability, AoA and subjective frequency. All the questionnaires were created using LimeSurvey (www.limesurvey.org). The participants were tested collectively in small groups (no more than eight individuals) in a large, quiet room. Each participant had a personal computer running the LimeSurvey questionnaire. Participants were randomly assigned to only one of the seven different rating tasks. The participants first gave their consent. Before each type of rating, full instructions with examples were given[3]. The 510 compounds were randomly presented across participants. Small breaks were allowed during the rating task.

To collect the ratings for lexeme meaning dominance, we closely followed the instructions provided by Juhasz et al. (2015). The participants were asked to assess the extent to which the meaning of the whole compound word was related more closely to the first lexeme or, in contrast, more closely to the second lexeme. A 10-point Likert scale was used, with 0 corresponding to the response "the meaning of the entire compound word is strictly related to the first lexeme" and 10 to "the meaning is strictly related to the second lexeme".

Semantic transparency, imageability, conceptual familiarity and sensory experience ratings were collected on the basis of the instructions used by Juhasz et al. (2015), except that we used 5-point Likert scales for all characteristics. As far as semantic transparency is concerned, the participants had to rate how related the two lexemes were to the overall meaning of the whole compound, with 1 corresponding to the response "the two words are not related at all to the meaning of the entire compound word", and 5 "the words are both completely related to the meaning of the entire compound word". To collect SER, the participants were told to indicate the degree of sensory experience that each compound word evoked for them, with 1 = "the word evokes no sensory experience" and 5 = "the word evokes a strong sensory experience". Imageability was rated by asking the participants to indicate how easily the concept denoted by the compound word made it possible to generate a mental image, with 1 = "the concept hardly evokes a mental image" to 5 = "the concept easily evokes a mental image". For the conceptual familiarity task, the participants were instructed to indicate how familiar the concept denoted by each compound word was for them (1 = "the concept is not familiar at all"; 5 = "the concept is very

---

[3] The following compounds were used as examples: *pommes de terre* (potatoes), *maréchal-ferrant* (blacksmith), *pèse-bébé* (baby scale) and *arme blanche* (white weapon = knife or sword).

familiar"). We followed the procedure described in Alario and Ferrand (1999) for collecting AoA norms. Five-point Likert scales with 3-year age bands were used, with "1" corresponding to "word acquired at 1–3 years old", and "5" corresponding to "word acquired at 13 years old or later". Finally, to evaluate the subjective frequency of compounds, the adults were told to indicate the degree to which they think they had read, heard or used them (1 = "I have never heard, read or used the word"; 5 = "I have very frequently heard, read or used the word").

Results

The Supplemental Material contains the database of 506 French compound words used in the present study (in csv and xls format; see below for the elimination of four compounds).

The compound words are listed alphabetically with their English translations. The database provides the mean ratings and standard deviations for the seven collected variables for each compound: lexeme meaning dominance, semantic transparency, sensory experience, conceptual familiarity, imageability, AoA and subjective frequency.

Data analyses

Several analyses were performed on the data and are reported in the following order. First, we describe the screening procedure that was used to identify potential outliers among the participants and the items. Second, we describe the reliabilities that were computed for the different collected norms. Third, descriptive statistics are reported together with the distributions of the norms. Fourth, we present the bivariate correlations and comment on a factor analysis that was conducted to analyze and summarize the correlational structure of the norms. Finally, we report multiple regression analyses that were performed in order to study the influence of the characteristics of the lexemes on compound ratings.

Screening of the data

Three participants were set apart: one in the imageability rating task, one in the conceptual familiarity rating task and, finally, one in the semantic transparency rating task[4]. These participants were eliminated because of the very low means of the correlations that were found between the estimations provided by them and the ratings obtained from the other

---

[4] More precisely, we computed the correlations between the ratings provided by each pair of participants, and then the by-participants' means of these correlations. In any given rating task, a participant was excluded if the mean of her/his correlations was three or more standard deviations below or above the mean of the correlation means.

participants in the respective rating tasks. Each compound word was evaluated by at least 28 participants. Four words were discarded ("appel d'offre" [invitation to tender], "appui-tête" [headrest], "belle-mère" [mother-in-law], "gratte-ciel" [skyscraper]) because they were inadvertently omitted or incorrectly spelled in certain rating tasks.

Reliability

The correlations between the by-item means obtained from the even and odd participants and the intraclass correlation coefficients (random effects of both participants and items—ICC(2, k) in Shrout and Fleiss's [1979] terminology) are reported in Table 1. With no values below .80, these coefficients suggest a high level of reliability for all the collected variables.

Descriptive statistics

Table 2 reports descriptive statistics for the different psycholinguistic variables and Fig. 1 shows the distributions of the different norms. As is generally observed for single words (e.g., Baayen, 2001; Muller, 1977), the objective frequency measures were very positively skewed. In line with Juhasz et al.'s (2015) study, most of the words had lexical frequencies below one per million: Among the 394 words for which objective frequencies were available, 381 (96.7%) had objective frequencies below 10 and only one (.3%) over 100. A positive skew of subjective frequency can also be observed, but is much smaller (Fig. 1).

With a relatively large negative skew (Fig. 1) and with both mean and median above 4, the conceptual familiarity ratings turned out to be relatively high, and they are indeed higher than those reported for French, Spanish and American English (the means were respectively 3.29, 3.12 and 3.06 on a 1–5-point scale) by Alario and Ferrand (Alario & Ferrand, 1999, see their Table 2, p. 534) for the Snodgrass and Vanderwart (1980) set of pictures of familiar objects. This characteristic

Table 1 Correlations between the by-items means obtained from the even and odd participants and the intraclass correlation coefficients (random effects of both participants and items—ICC(2, k) in Shrout and Fleiss's (1979) terminology)

	R(even,odd)	ICC
Age of acquisition	.95	.96
Conceptual familiarity	.87	.89
Subjective frequency	.94	.94
Lexeme meaning dominance	.81	.87
Semantic transparency	.93	.96
Imageability	.91	.94
Sensory experience rating	.83	.87

Table 2 Descriptive statistics for the subjective norms and other objective variables

	AoA	Fam	Imag	LMD	Trans	SER	S-Freq	Length	O-Freq	
									Subtitle	Book
Mean	3.35	4.01	3.52	5.20	3.39	2.45	2.95	11.77	2.74	2.76
SD	.77	.67	.89	.98	.99	.67	.81	2.33	19.63	11.71
Q1	2.74	3.60	2.80	4.67	2.53	1.94	2.32	10.00	.07	.14
Median	3.45	4.17	3.60	5.30	3.50	2.35	2.90	12.00	.30	.54
Q2	3.94	4.53	4.27	5.83	4.27	2.90	3.58	13.00	.87	1.49
Skew	−.29	−.89	−.27	−.43	−.30	.48	.28	.39	15.95	9.33
Kurtosis	−.68	.24	−.97	.30	−1.03	−.34	−.73	.45	283.03	100.99
Min	1.52	1.70	1.60	2.10	1.13	1.13	1.23	6.00	.00	.00
Max	4.90	5.00	4.97	7.50	4.97	4.39	4.97	21.00	360.17	158.38

Notes. O-Freq = objective subtitle and book frequencies; S-Freq. = subjective frequency; Fam = conceptual familiarity; AoA = age of acquisition; Trans = semantic transparency; LMD = lexical meaning dominance; Imag = imageability; SER = sensory experience rating. SD = standard deviation; Q1 = 25th percentile; Q3 = 75th percentile. Subtitle and book frequencies were taken from Lexique (New et al., 2004) and were available for only 394 items

suggests that the ratings given for the familiarity of the concepts are dependent on the format of presentation (words versus pictures).

Lexical meaning dominance and numbers of letters had nearly symmetrical distributions (Fig. 1). The other norms were not strongly skewed and their main characteristic resides in a negative excess kurtosis, with thinned tails and relatively flat distributions between the tails. In contrast to conceptual familiarity, the characteristics of the AoA and imageability variables were consistent with what has been found for single words in previous databases (Alario & Ferrand, 1999; Bonin, Méot, et al. 2003a; Bonin, Peereman, et al., 2003b). Indeed, compound words were estimated to be learned late in life, and to be less imageable, than single words. The differences were tenuous when compared to subjective frequency as reported by Bonin, Méot, et al., (2003a).

Bivariate correlations and multiple regressions with behavioral variables as dependent variables

Pearson correlations (see Table 3) were computed between the mean ratings for each item on all of the rating measures as well as with objective measures.

Table 4 shows the loadings obtained from a varimax rotation computed on the three components associated with eigenvalues beyond 1 in the corresponding principal component analysis (PCA; it is worth noting that the pattern of findings was nearly the same without the inclusion of objective frequency). Subjective frequency, conceptual familiarity, imageability and SER were essentially expressed on the first component and they were very positively correlated, i.e., compounds that were rated highly on one variable also tended to be rated highly on the other variables. There were also strong

relationships between this set of variables and AoA, with later-acquired compounds being judged as less frequent, conceptually familiar and imageable and as being associated with fewer sensory experiences than early-acquired compounds. With noticeable correlations with conceptual familiarity, subjective frequency and AoA, objective frequency and semantic transparency were also partially expressed on this first factor, with more frequent and transparent words rated as being somewhat more familiar, subjectively frequent and earlier acquired (and also more imageable in the case of semantic transparency only) than less frequent and more opaque words. Length was poorly correlated with all the other variables, and this variable was expressed on the second factor, in part linked with objective frequency and semantic transparency by means of weak negative correlations, i.e., longer words tended to be associated with lower frequency values and higher transparency ratings. Finally, the third factor essentially expressed lexeme meaning dominance, which was nearly uncorrelated with all the other variables. The only significant correlation was with imageability, but with an opposite sign to that found by Juhasz et al. (2015).

Taken overall, the pattern of correlations was very similar to that obtained by Juhasz et al. (2015). The only noticeable difference concerns the correlation between semantic transparency and SER which, despite being of the same sign, was clearly lower in the current study. It is important to note that the correlations between objective lexical frequency and the other variables reported by Juhasz et al. (2015) were always lower than the correlations reported here for French compounds. However, it seems that the authors did not log-transform objective frequency before performing the computations, and that is the reason why we also report the correlations with raw objective frequency in parentheses in
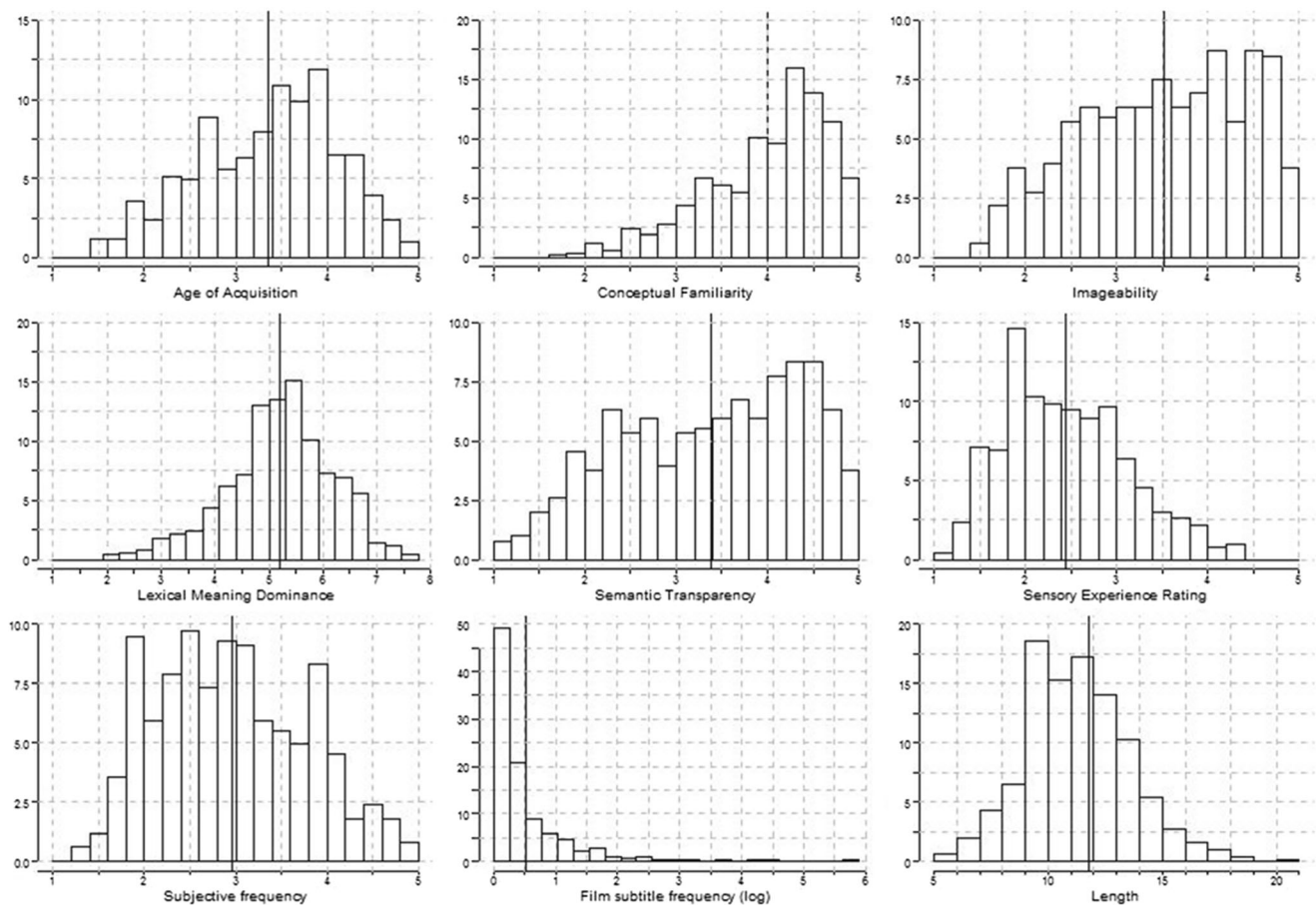
Fig. 1 Distributions of compounds as a function of the different variables (vertical lines = means of the variables)

Table 3. Indeed, these latter correlations are approximately similar to those obtained by Juhasz et al. (2015).

In order to examine whether the relationships described above persisted when other dimensions were controlled for, we included each psycholinguistic variable in turn as the dependent variable in a multiple regression analysis and used the other variables as independent variables. (It is worth mentioning that insofar as a PCA—with or without varimax rotation—depends solely on bivariate correlations, it was not suitable for achieving this aim. The same objection could have been made if a clustering method had been used.) Given the high correlation between subjective frequency and conceptual

familiarity, only subjective frequency was retained in the analyses. To simplify the comparisons, the variables were all transformed into Z-scores. Non-linearities were introduced by including restricted cubic splines with a maximum of six knots for some independent variables[5]. In order to reduce potential overfitting[6], we decided to use a bootstrap approach which was suggested to us by Baayen (2008, p. 212). In a first step, 1000 bootstrap samples were selected. In each sample, each subjective norm was taken in turn as a dependent variable (DV) whereas the other norms were included as independent variables (IV). For each DV, the model with no nonlinearities was enriched by including, in turn, splines with three to six knots for each IV. The percentage of results significant at .05 over all bootstrap samples was then computed for each IV both overall and for its nonlinear part. Nonlinearities were retained for the IV whenever (1) these percentages were the highest and (2) both percentages were higher than 95% for at least one number of knots. Given that beyond the first number of knots that satisfied these criteria, the numbers of subsequent knots led to very similar percentages of significant results, we then computed, over all bootstrap samples, the mean of the R^2 using three to six knots for the retained IV. The retained number of knots was the one after which adding more knots did

---

[5] It is worth noting that nonlinearities usually exist independently of the "statistical control" operated by multiple regression and can generally be identified by considering pairs of plotted variable (e.g. Baayen, 2008, p. 191), possibly together with fitted values obtained with bivariate regressions allowing for nonlinearities (e.g. Loess regression). Bivariate linear correlation and simple linear regression coefficients are not suited to studying nonlinearities.

[6] In a previous version of this paper, we used a forward procedure which was similar to that described in Bonin, Méot, and Bugaiska's (2018) study. This procedure gave rise to smoothings that were sometimes specific to certain restricted parts of the independent variables, more particularly for the semantic transparency variable. It is worth noting, however, that the nonlinear effects found with the bootstrap approach follow roughly the same patterns as those found with the initial procedure.

Table 3 Bivariate correlations between the measured variables

	O-Freq	S-Freq	Fam	AoA	Trans	LMD	Imag	SER
Length	−.186*** (−.057)	−.040	−.007	.073	.134**	.017	.091*	.025
O-Freq		.521*** (.247)	.378*** (.149)	−.394*** (−.213)	−.011 (.005)	.087† (−.036)	.155*** (.031)	.108** (.008)
S-Freq			.869***	−.611***	.303***	.017	.523***	.415***
Fam				−.668***	.379***	−.021	.670***	.525***
AoA					−.297***	.031	−.738***	−.579***
Tran						.076†	.345***	.147***
LMD							−.101*	−.087†
Imag								.682***

Notes. O-Freq = objective subtitle frequency (in log+1); S-Freq. = subjective frequency; Fam = conceptual familiarity; AoA = age of acquisition; Trans = semantic transparency; LMD = lexical meaning dominance; Imag = imageability; SER = sensory experience rating. Correlations with O-Freq are given for 394 words. In brackets, correlations between raw frequencies. ***$p < .001$; **$p < .01$; *$p < .05$; † $p < .10$

not add more than 1% of additional explained variance (as a mean over all bootstrap samples). In a second step that was limited to the DVs for which one IV including nonlinearities was retained, the same procedure was repeated in order to determine whether such terms should be included for any of the other IVs (the baseline model included nonlinear terms for the IV retained in the first step and only linear terms for the other IVs). However, no IV satisfied the criteria defined above. As a result, no more nonlinearities were included.

The results of these analyses are reported from two perspectives: (1) Table 5 shows the squares of the semi-partial correlations, i.e., estimations of the additional percentages of explained variance when the independent variables are added in the model while the other independent variables are already included in the model. (2) Significant partial effects of the independent variables, that is to say the effects of the independent variables measured at constant levels of the other independent variables, are depicted in Fig. 2. The aim of this figure is to illustrate the comparison of the relationships that exist between pairs of subjective norms when the other dimensions are controlled for, with the associations appearing in the form of bivariate correlations, and also to illustrate the patterns of nonlinear relationships revealed by the inclusion of nonlinear terms.

For each IV, the added percentage of explained variance when it was included while other IVs were already in the model is shown in the first line. If nonlinear terms were included, a second line gives the same information for nonlinear terms alone together with, in brackets, the number of knots used. The analyses were run using the reduced set of (394) compounds for which subtitle frequencies were available in Lexique (New et al., 2004). ***$p < .001$; **$p < .01$; *$p < .05$

With approximately 70% of explained variance, imageability was predicted to a large extent by the other dimensions, in particular SER, with compounds rated as easily experienced by the senses being more imageable, and AoA, with later-acquired words being less imageable. The increase in imageability ratings with higher SER values was, however, somewhat reduced after the SER mean.

More transparent and longer compounds were also more imageable and the reverse was true for objective high-frequency compounds. There were some discrepancies between the bivariate correlations, and more particularly for the subjective frequency variable. In effect, in spite of a large positive correlation, there was no reliable effect when the other dimensions were controlled for. In addition, the effect of objective lexical frequency was reversed, with more frequent

Table 4 Loadings obtained in the varimax rotation

Component	Length	O-Freq	S-Freq	Fam	AoA	Trans	LMD	Imag	SER
1 (42.55%)		0.40	0.81	0.89	−0.86	0.44		0.86	0.74
2 (14.63%)	0.74	−0.65	−0.23			0.48			
3 (12.81%)		0.35	0.26			0.36	0.83		−0.30

Notes. O-Freq = objective subtitle frequency (in log+1); S-Freq = subjective frequency; Fam = conceptual familiarity; AoA = age of acquisition; Trans = semantic transparency; LMD = lexical meaning dominance; Imag = imageability; SER = sensory experience rating. Explained variance is given in brackets after each component number. Loadings under .2 are not shown

Table 5 Multiple regressions including one norm as DV and the other norms as IVs

	S-Freq	AoA	Trans	LMD	Imag	SER
R^2	.532 (.554)	.651	.209 (.249)	.042	.673 (.697)	.498
Length	.000	.006**	.008*	.002	.005*	.000
O-Freq	.113***	.021***	.019***	.002	.005*	.008*
S-Freq	-	.021***	.045***	.005	.001	.013**
AoA	.020***	-	.044***	.000	.104***	.005*
non linear			*.039*** (5)*			
Tran	.029***	.002	-	.005	.006**	.019***
LMD	.003	.000	.003	-	.001	.003
Imag	.024***	.124***	.018**	.006	-	.134***
non linear	*.022*** (4)*				-	
SER	.011**	.004*	.026***	.005	.112***	-
non linear			-	-	*.024*** (3)*	-

Notes. O-Freq = objective subtitle frequency; S-Freq = subjective frequency; Fam = conceptual familiarity; AoA = age of acquisition; Trans = semantic transparency; LMD = lexical meaning dominance; Imag = imageability; SER = sensory experience rating. For each IV, the added percentage of explained variance when it was included while other IVs were already in the model is shown in the first line. If nonlinear terms were included, a second line gives the same information for nonlinear terms alone together with, in brackets, the number of knots used. The analyses were run using the reduced set of (394) compounds for which subtitle frequencies were available in Lexique (New et al., 2004). ***$p < .001$; **$p < .01$; *$p < .05$

$R^2 = R^2$ with no nonlinearities and, in brackets, with nonlinearities included.

words being less imageable when other variables were entered in the equation. Given that these discrepancies were still observed when one of the two frequency measures was excluded from the equation, they cannot be accounted for by the relationship between the two frequency measures. This might have to do with the relationship between subjective frequency and AoA. Indeed, excluding AoA from the regression equation led to a large positive effect of subjective frequency. The results of the multiple regression analysis for SER mimic in part those that were found for imageability (even though SER is less well explained by the other variables than imageability): The more imageable the compounds were rated as being, the more they were rated as easily experienced by the senses. Late-acquired and more objectively frequent compounds were rated as arousing less sensory experience.

Contrary to what occurred for imageability, the positive effect of subjective frequency found in the bivariate correlation analysis was still observed when the other independent variables were controlled for. Finally, SER were lower for more transparent words in contrast to what was observed in the bivariate correlation analysis.

The variance in the AoA ratings was explained by the other variables at a level similar to that of imageability. The effects were in the same directions as those obtained in the bivariate correlations: Higher imageability and SER ratings as well as higher (objective and subjective) frequencies were associated

Table 6 Regressions including left (L) and right (R) word ratings

	S-Freq	AoA	Trans	LMD	Imag	SER
n	137	134	360	360	137	126
Adj. R2	.317 *(.262)*	.284 *(.195)*	*(.02)*	*(.057)*	.344 *(.263)*	.260 *(.172)*
Length	.047 *(.082)*	−.054 *(−.137†)*	.130*	.021	.176* *(.278***)*	.083 *(.154†)*
Freq	.466*** *(.449***)*	−.428*** *(−.403***)*	−.037	.162**	.263*** *(.272***)*	.114 *(.205*)*
L-Freq	−.091 *(−.076)*	.218* *(.202*)*	.139*	.178***	−.263*** *(−.338***)*	−.286*** *(−.363***)*
R-Freq	−.003 *(.220**)*	.114 *(−.04)*	.060	−.138*	.034 *(−.024)*	.109 *(.104)*
L-Rating	.011	.110			.219**	.135†
R- Rating	.330***	.314***			.193*	.297***

Notes. O-Freq = objective subtitle frequency (in log+1); S-Freq = subjective frequency; Fam = conceptual familiarity; AoA = age of acquisition; Trans = semantic transparency; LMD = lexical meaning dominance; Imag = imageability; SER = sensory experience rating. Upper part: *Adj. R2*: Adjusted R square in the model including ratings for the compounds' constituents as well as in the model excluding them (in brackets). Following lines: for each IV, estimated betas in the two models (in brackets for the model excluding constituents' ratings). Lower part: estimated betas for the ratings of the constituents.

***$p < .001$; ** $p < .01$; *$p < .05$; † $p < .10$
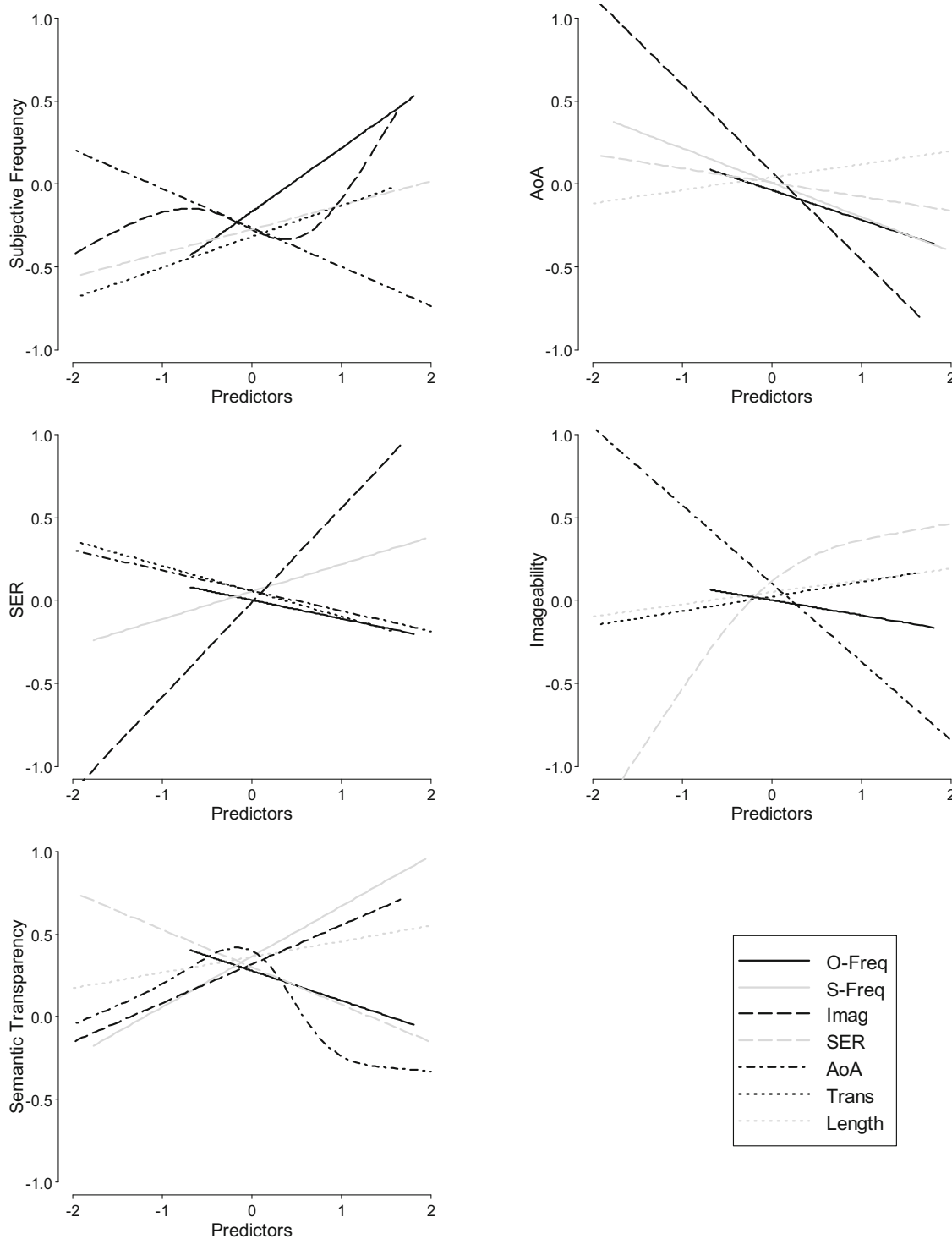
Fig. 2 Partial effects in the regression analyses including one variable as the DV and the other variables as independent variables (AoA = age of acquisition; Trans = semantic transparency; Imag = imageability; S-Freq = subjective frequency; O-Freq = objective frequency; LMD = lexical meaning dominance; SER = sensory experience ratings)

to a greater extent with earlier acquired compounds. Imageability was the most explanatory variable, accounting for about 12% of explained variance. There were, however, two exceptions. First of all, semantic transparency failed to reach significance whereas its bivariate correlation was

relatively large, and second, a significant effect of length was found, with longer compounds being acquired later.

As far as subjective frequency is concerned, a large proportion of the variance was accounted for by the other variables, with the effects having directions similar to those shown in the

bivariate correlation analysis. Objective frequency was the most explanatory variable, with more objectively frequent compounds being rated as more frequently encountered. AoA and semantic transparency also accounted for a large portion of the variance: More transparent compounds were rated as being encountered more often, and the reverse was true for late-acquired words. With a less strong effect, compounds with high sensory experience ratings were judged to be encountered more often. This was also the case of imageability, but its effect was not monotonous, with subjective frequency ratings increasing roughly before −1 sd and after +1 sd of the ratings of imageability and the reverse for intermediate ratings.

Compared to the other psycholinguistic variables, differences in semantic transparency estimations were far less well explained by the other dimensions. They increased with higher subjective frequency and imageability ratings and, to a lesser extent, with longer compounds (and the opposite for objective lexical frequency and SER). Finally, the link with AoA was more complex, with increasing transparency values located before the mean AoA and decreasing values located after the mean AoA.

Finally, lexical meaning dominance was very poorly explained by the other variables, with no predictors reaching significance. Conversely, lexical meaning dominance was never significant when it was introduced as IV in the analyses.

Influence of lexeme characteristics on compound ratings

In order to compare our results with those obtained by Juhasz et al. (2015) concerning the influence of the characteristics of the constituents of the compounds, each compound rating was used as a dependent variable in a linear regression which included the length of the compound, its subtitle frequency, the subtitle frequencies of its constituents (subtitle frequencies were all taken from Lexique: New et al., 2004) and the ratings of the left and right constituents, hereafter L-words and R-words, respectively. Because French has fewer databases of psycholinguistic norms than are available for English, and because all the French databases have—numerically speaking—substantially fewer rated words than in English, the ratings for L-words and R-words were available for only a subset of the compound words. As a result, we decided to combine different French databases in order to maximize the numbers of ratings. Before doing this, we standardized the ratings within each database and checked whether the results obtained using this transformation were approximately the same as those resulting from the use of only the database containing the greatest number of words.

The AoA ratings for the constituent words were obtained from Ferrand et al.'s (2008) study ($N = 114$), from Alario and Ferrand's (1999) study ($N = 12$) and from Bonin, Peereman

et al.'s (2003) study ($N = 8$). Imageability ratings were taken from Bonin et al.'s (2011) study ($N = 114$) and from Bonin, Méot, et al.'s (2003) study ($N = 23$). SER ratings were obtained from Bonin et al. (2015). We also took into account the subjective frequency ratings of the constituents. These were obtained from Ferrand et al.'s (2008) study ($N = 114$) and from Bonin, Méot et al.'s (2003) study ($N = 23$).

The analyses were run in two steps. First of all, the analyses were performed with a block containing the independent variables common to all ratings, with length and objective word frequency included. Second, they were run with a block including the ratings of the constituents.

The percentages of explained variance were roughly comparable to those found by Juhasz et al. (2015) for all dependent variables except subjective frequency, for which it was a little lower in the current study when the ratings of the constituents were not included as they were in Juhasz et al. (2015). Even though they were all significant at .001, the increase in the percentages of explained variance when the lexeme ratings were included in addition to other characteristics were lower in the present study than in the Juhasz et al. (2015) study (AoA: .089 versus .134; imageability: .081 versus .239 ; SER: .088 versus .186).

The compound frequency effects were also similar to those observed by Juhasz et al. (2015), albeit with a nonsignificant effect in the case of SER, for which the estimated beta was nearly half that reported by the latter authors. For both subjective frequency and AoA, these were the largest effects, with more objectively frequent compounds being rated as more frequently encountered and acquired earlier in life.

The negative effects of the frequency of L-words on imageability and SER were also replicated, but they were stronger in the present study than in Juhasz et al.'s (2015) study. For imageability, there was an absolute effect which was as large as that of compound frequency. By contrast, the effect of the frequency of L-words on AoA was positive but less strong than in Juhasz et al. (2015). Unlike in this latter study, the frequency of L-words also had significant positive effects on both semantic transparency and lexical meaning dominance. However, these effects turned out to be relatively weak. Finally, no reliable effects of the frequency of R-words were found, except for a small negative effect on lexical meaning dominance.

As far as the effects of the ratings of the L- and R-words are concerned, our findings were less consistent with those obtained by Juhasz et al. (2015). In their study, both the ratings of L-words and R-words had significant effects on compound ratings, with the effects of the first being larger than those of the second. In contrast, with the exception of imageability, the effects of the L-word ratings failed to reach significance in the present study. Moreover, with the exception of imageability, the effects of the R-word ratings were also somewhat larger in the present study. It is important to note that for both subjective

frequency and AoA, this property was also observed in bivariate correlations between L-word and R-word ratings and the compound ratings (L-words: $r = -.02$ and $r = .043$; R-words: $r = .36$ and $r = .31$), whereas this was less the case for SER (L-words: $r = .23$ and R-words: $r = .42$) and not at all the case for imageability (L-words: $r = .39$ and R-words: $r = .39$).

Discussion

Our main motivation in the present research was to provide psycholinguistic norms for a set of French compound words. The compounds were normed on seven psycholinguistic variables: lexeme meaning dominance, semantic transparency, sensory experience, conceptual familiarity, imageability, age of acquisition (AoA) and subjective frequency.

Let us summarize the main findings. First of all, and crucially, the reliabilities were high for all the collected variables. Second, and interestingly, the pattern of bivariate correlations among the norms was very similar to that reported by Juhasz et al. (2015), which suggests that the underlying structure of the norms is similar in both English and French. The important aspects of note among the correlations between the variables include the following: (1) Compounds that were given high ratings on subjective frequency, familiarity, imageability or SER—or that were estimated to be acquired early in life— were given higher ratings on the other dimensions. The relations of these variables with objective frequency were in the same direction but weaker. (2) Semantic transparency was also judged to be higher for higher ratings on the variables listed above (e.g., subjective frequency, familiarity, imageability). The same was true for early acquired words, but the relationships were more tenuous. Semantic transparency was also the variable which was the most strongly related to orthographic length with, however, only a small correlation of .134. (3) Lexical meaning dominance ratings were only very weakly related to other psycholinguistic variables.

Taken overall, the relationships described above were also observed when other dimensions were controlled for using multiple regression analyses. First of all, except for AoA for which the opposite was found, imageability was positively related to all the other psycholinguistic norms, and more strongly to SER and AoA. In spite of the fact that SER was generally less explanatory than imageability, the relations were in the same direction as those observed with imageability. One exception was semantic transparency, for which the more the compounds were judged as being easily experienced by the senses, the less transparent they were rated to be. Second, compared to early-acquired compounds (e.g., *petit déjeuner* [breakfast], *pique-nique* [picnic]), late-acquired compounds (e.g., *trop-perçu* [overpayment], *arc-boutant* [flying buttress]) were rated as less imageable, less easily experienced by the senses and less frequent. Also, both early- and late-acquired compounds were judged to be less transparent than compounds having intermediate age of acquisition ratings. Third, objectively high-frequency compounds (e.g., *aujourd'hui* [today], *grand-père* [grandfather], *rendez-vous* [appointment]) were estimated to be more frequently encountered and acquired later than compounds that are objectively of low-frequency (e.g., *abaisse-langue* [tongue depressor], *trop-perçu* [overpayment]). Objectively high-frequency compounds were also judged to be less imageable and easily experienced by the senses, and less semantically transparent, than objectively low-frequency compounds. Finally, more semantically transparent compounds were rated as being more frequently encountered and imageable, but less easily experienced by the senses whereas longer compounds were judged more imageable and transparent, and as having been acquired later than shorter compounds.

The multiple regression analyses that were performed to examine the influence of lexeme characteristics on compound ratings revealed that, except for imageability, there were no reliable effects of the L-word ratings which contrast with Juhasz et al.'s results (2015) in which reliable and large effects of L-word ratings were found for AoA and SER. As far as R-word ratings are concerned, except for imageability, the impact of R-word ratings was also larger in the present study than in Juhasz et al.'s (2015). However, including lexeme ratings added a relatively high proportion of explained variance for all norms. Taken as a whole, this pattern of findings accords with the idea that compound words are decomposed in order to perform rating.

Different characteristics of compounds have been found to modulate the speed and accuracy of the processing of compounds in several word reading or word production tasks, thus suggesting that the constituents of compounds are activated to some degree and that they influence the processes underpinning language processing. As reviewed in the Introduction, certain psycholinguistic characteristics of compounds have been especially well-investigated (e.g., lexical frequency, semantic transparency), whereas other characteristics have given rise to a limited number of studies (AoA, conceptual familiarity). Thus, the influence of the latter variables should be addressed in future studies in order to determine their true influence. Therefore, we hope that the availability of these norms for the French language will stimulate research on how compound words are processed in word reading and in both spoken and written word production. For instance, it would be interesting to design a megastudy including a large number of French compounds to determine the determinants of recognition performance (in lexical decision, in word reading or in perceptual identification). Indeed, the megastudy approach is relatively recent for investigating lexical processing and it is very useful in that it permits the investigation of a large number of variables (and sometimes novel variables) and their relationships (including complex relationships such

as nonlinear ones) on a large number of items (Balota et al., 2012; Yap & Balota, 2015). In the future, having large databases on compound word processing speed and accuracy in French should also make it possible to investigate individual differences in compound processing (e.g., processing differences when compounds are read by high- versus low-skilled readers; see Yap et al., 2012, and Lim et al., 2020, for examples of this approach in visual word recognition). As far as compounds are concerned, megastudies have been conducted in English (Kim et al., 2019) and in Chinese (Tse & Yap, 2018). These have helped identify the characteristics that play a significant role in compound word reading and those that appear to be of less importance. For instance, in Chinese, Tse and Yap (2018) found that orthographic variables accounted for the largest part of variance in lexical decision, followed by semantic variables, and finally by phonological variables. We plan to use our norms to design such a study in the future.

Compared to certain more recent studies that have rated compounds on impressive numbers of items (e.g., Gagné et al. (2019) collected meaning predictability judgments for 8,304 compounds and Kim et al. (2019) collected semantic transparency scores for 2,861 compound words), our database is somewhat smaller in terms of the number of compounds that are rated. However, our compounds have been normed on seven psycholinguistic variables, and not only on one or two variables (e.g., imageability, semantic transparency), a fact which will be very useful for many experimental designs. A potential limitation of the current work is that the norms were collected from French-speaking adults who live in France. We cannot exclude the possibility that the norms may not be entirely suitable for use with adults speaking French in Belgium, Switzerland, or parts of Africa. It would be interesting to collect norms on our compounds with such French speakers and to compare them with the norms reported here. It is worth mentioning in passing that norms for compound words are less frequent in languages other than English and, as has been the case for AoA or imageability norms, it is possible that norms for compounds will become available for other alphabetic languages as well as for a larger number of non-alphabetic languages. Finally, as has been the case for other pools of words (e.g., Ferrand et al., 2008), the current norms will be complemented in the future by other types of norms, such as association norms (see Schulte im Walde & Borgwaldt, 2015, for a study of this kind in German), valence and arousal (Kuperman, 2013) or "word prevalence" norms, which have recently been collected for single words (Brysbaert et al., 2019).

To conclude, we hope to have convinced readers that norms on compounds such as those provided here are indispensable for achieving a better understanding of how compounds are comprehended or verbally produced.

Appendix

Table 7 List of publications in which one or more psycholinguistic variables were studied for compound words

Source	Year of publication	Language	Lexical frequency	Length	Semantic transparency	Familiarity/ Subjective frequency	AoA	Lexeme dominance	Imageability	SER
Taft & Forster	1976	English	✓	✓c						
Sandra	1990	Dutch			✓					
Zhou & Marslen-Wilson	1994	Chinese	✓							
Zwitserlood	1994	Dutch			✓					
Pratarelli	1995	English			✓					
Inhoff, Bruhl, & Schwartz	1996	English	✓							
Roelofs	1996	Dutch	✓							
Hyönä & Pollatsek	1998	Finnish	✓	✓ ✓c		✓c				
Tamaoka & Hatsuzuka	1998	Japanese		✓						
Jarema, Busson, Nikolova, Tsapkini & Libben	1999	French and Bulgarian			✓					
Pollatsek, Hyönä, & Bertram	2000	Finnish	✓							
Badecker	2001	English	✓							
Bertram & Hyönä	2003	Finnish	✓	✓						
Fang	2003	Chinese	✓	✓						
Juhasz, Starr, Inhoff, & Placke	2003	English	✓	✓c						
Andrews, Miller & Rayner	2004	English	✓	✓c						
Hyönä & Bertram	2004	Finnish	✓	✓ ✓c						

Table 7 (continued)

Source	Year of publication	Language	Lexical frequency	Length	Semantic transparency	Familiarity/ Subjective frequency	AoA	Lexeme dominance	Imageability	SER
Hyönä, Bertram, & Pollatsek	2004	Finnish	✓							
Koester, Gunter, Wagner, & Friederici	2004	German	✓c							
Bien, Levelt, & Baayen	2005	Dutch	✓							
Christianson, Johnson, & Rayner	2005	English	✓c							
Pollatsek & Hyönä	2005	Finnish	✓		✓					
Chen & Chen	2006	Chinese	✓							
Ogawa & Saito	2006	Japanese	✓c							
Duñabeitia, Perea, & Carreiras	2007	Spanish and Basque	✓							
Fiorentino & Poeppel	2007	English	✓	✓c						
Bi, Han, & Shu	2007	Chinese	✓							
Koester, Gunter, & Wagner	2007	German	✓c	✓c	✓					
Nicoladis & Krott	2007	French (children)	✓							
Tamaoka	2007	Japanese	✓c			✓c				
Frisson, Niswander-Klement, & Pollatsek	2008	English	✓c	✓ ✓c	✓					
Inhoff, Starr, Solomon, & Placke	2008	English	✓						✓	
Janssen, Bi, & Caramazza	2008	English and Chinese	✓	✓c		✓				
Juhasz	2008	English	✓	✓		✓				
Juhasz, White, Liversedge, & Rayner	2008	English		✓						
Pollatsek, Slattery, & Juhasz	2008	English	✓	✓ ✓c						
Koester & Schiller	2008	Dutch	✓c	✓c	✓					
Duñabeitia, Laka, Perea, & Carreiras	2009a	Basque:	✓c	✓c						
Duñabeitia, Marín, Avilés, Perea, & Carreiras	2009b	Spanish			✓c					
Gagné & Spalding	2009	English	✓							
Juhasz, Pollatsek, Hyönä, Drieghe, & Rayner	2009	English	✓							
Mok	2009	Chinese	✓		✓					
Vergara-Martínez, Duñabeitia, Laka, & Carreiras	2009	Basque	✓							
Desrochers, Liceras, Fernández-Fuertes, & Thompson	2010	Spanish	✓			✓				
Drieghe, Pollatsek, Juhasz, & Rayner	2010	English	✓	✓						
Häikiö, Bertram & Hyönä	2010	Finnish	✓	✓c						
Kaakinen & Hyönä	2010	Finnish	✓	✓						
Huang, Lee, Tsai, & Tzeng	2011	Chinese	✓			✓c				
Hyönä & Bertram	2011	Finnish	✓	✓						
Janssen, Pajitas, & Caramazza	2011	English	✓			✓				
Juhasz & Berkowitz	2011	English	✓c	✓ ✓c						
Ko, Wang, & Kim	2011	Korean and English	✓		✓c					
Koester & Schiller	2011	Dutch	✓c	✓c	✓					
Lüttmann, Bölte, Böhl, & Zwitserlood	2011	German			✓					
Pollatsek, Bertram, & Hyönä	2011	Finnish	✓	✓c						
	2011	Italian	✓c			✓		✓c	✓c	

Table 7 (continued)

Source	Year of publication	Language	Lexical frequency	Length	Semantic transparency	Familiarity/ Subjective frequency	AoA	Lexeme dominance	Imageability	SER
Semenza, Arcara, Facchini, Meneghello, Ferraro, Passarini, Pilosio, Vigato, & Mondini										
Fehringer	2012	English			✓					
Forgács, Bohrn, Baudewig, Hofmann, Pléh, & Jacobs	2012	German	✓c	✓c					✓c	
Juhasz	2012	English	✓	✓c	✓c					
Marelli & Luzzatti	2012	Italian	✓	✓	✓			✓ ?		
Zhang, Anderson, Wang, Packard, Wu, Tang, & Ke	2012	Chinese and English	✓							
Bronk, Zwitserlood, & Bölte	2013	German	✓	✓c	✓					
Crepaldi, Rastle, Davis, & Lupker	2013	English	✓c	✓c						
Cui, Yan, Bai, Hyönä, Wang, & Liversedge	2013	Chinese	✓	✓c						
Kuperman	2013	English	✓c	✓c					✓	✓
MacGregor & Shtyrov	2013	English	✓		✓					
Verdonschot, La Heij, Tamaoka, Kiyama, You, & Schiller	2013	Japanese	✓c							
Arcara, Semenza, & Bambini	2014	Italian	✓	✓						
Fiorentino, Naito-Billen, Bost, & Fund-Reznicek	2014	English	✓c							
Janssen, Pajitas, & Caramazza	2014	English	✓							
Marelli, Zonca, Contardi, & Luzzatti	2014	Italian	✓	✓c	✓		✓	✓c		✓✓c
Nakayama, Sears, Hino, & Lupker	2014	Japanese	✓							
Bertram, Tønnessen, Strömqvist, Hyönä, & Niemi	2015	Finnish	✓	✓						
Falkauskas & Kuperman	2015	English	✓	✓						
Juhasz, Lai, & Woodcock	2015	English	✓			✓	✓	✓	✓	✓
Pham & Baayen	2015	Vietnamese	✓							
Dasgupta, Sinha, & Basu	2016	Bangla			✓					
Blazej & Cohen-Goldberg	2016	English	✓							
Mankin, Thompson, Branigan, & Simner	2016	English	✓		✓					
Gagné & Spalding	2016a, b	English			✓					
Lorenz & Zwitserlood	2016	German	✓c	✓c	✓					
Reinhart, Schunck, Schaadt, Adams, Simon, & Kerkhoff	2016	German	✓	✓						
Cui, Häikiö, Zhang, Zheng, & Hyönä	2017	Chinese	✓	✓						
Li, Jiang, & Gor	2017	English	✓c							
Liang, Blythe, Bai, Yan, Li, Zang, & Liversedge	2017	Chinese	✓							
Tse, Yap, Chan, Sze, Shaoul, & Lin	2017	Chinese	✓		✓					
Wang, Lu, He, Zhang, & Zhang	2017	Chinese	✓							
Juhasz	2018	English	✓	✓		✓	✓	✓	✓	✓
Schmidtke, Van Dyke, & Kuperman	2018c	English	✓		✓					
	2018	English	✓							

Table 7 (continued)

Source	Year of publication	Language	Lexical frequency	Length	Semantic transparency	Familiarity/ Subjective frequency	AoA	Lexeme dominance	Imageability	SER
Schmidtke, Gagné, Kuperman, Spalding, & Tucker										
Schmidtke, Gagné, Kuperman, & Spalding	2018	English	✓							
Shen, Li, & Tong	2018b	Chinese	✓c							
Shen, Li & Pollatsek	2018a	Chinese	✓c							
Tse & Yap	2018	Chinese	✓		✓					
Afonso & Álvarez	2019	Spanish	✓							
Davis, Libben, & Segalowitz	2019	English	✓		✓					
Gagné, Spalding, & Schmidtke	2019	English	✓		✓					
Güemes, Gattei, & Wainselboim	2019	Spanish	✓							
Günther & Marelli	2019	English	✓		✓					
Hasenäcker & Schroeder	2019	German	✓							
Kim, Yap, & Goh	2019	English	✓c	✓c	✓					
Kush, Dillon, Eik, & Staub	2019	Norwegian	✓							
Schmidtke & Kuperman	2019	English	✓							
Elsherif, Catling, & Frisson	2020	English	✓		✓		✓		✓	
Günther, Marelli, & Bölte	2020	German	✓		✓					
Lim, Yap, & Tse	2020	Chinese	✓		✓					

Note. c = control variable

References

Afonso, O., & Álvarez, C. J. (2019). Constituent frequency effects in the written production of Spanish compound words. *Memory & Cognition, 47,* 1284–1296. **(Table, appendix)**

Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers, 31,* 531-552.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16,* 285-311. **(Table, appendix)**

Arcara, G., Semenza, C., & Bambini, V. (2014). Word structure and decomposition effects in reading. *Cognitive Neuropsychology, 31,* 184–218. **(Table, appendix)**

Baayen, R. H. (2001). *Word frequency distributions*. Dordrecht: Kluwer.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.

Badecker, W. (2001). Lexical composition and the production of compounds: Evidence from errors in naming. *Language and Cognitive Processes, 16,* 337–366. **(Table, appendix)**

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition, 29,* 639-647.

Balota, D. A., Yap, M. J., Hutchinson, K. A., & Cortese, M. J. (2012). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Eds.), *Visual word recognition volume 1: Models and methods* (pp. 90– 115). Psychology Press.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language, 48,* 615-634. **(Table, appendix)**

Bertram, R., Tønnessen, F. E., Strömqvist, S., Hyönä, J., & Niemi, P. (2015). Cascaded processing in written compound word production. *Frontiers in Human Neuroscience, 9:*207. **(Table, appendix)**

Bi, Y., Han, Z., & Shu, H. (2007). Compound frequency effect in word production: Evidence from anomia. *Brain and Language, 103,* 8-249. **(Table, appendix)**

Bien, H., Levelt, W. J. M., & Baayen, R. H. (2005). Frequency effects in compound production. *PNAS, 102,* 17876–17881. **(Table, appendix)**

Blazej, L. J., & Cohen-Goldberg, A. M. (2016). Multicolored words: Uncovering the relationship between reading mechanisms and synesthesia. *Cortex, 75,* 160-179. **(Table, appendix)**

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and language, 50,* 456-476.

Bonin, P., Méot, A., & Bugaiska, A. (2018a). Concreteness norms for 1, 659 French words: Relationships with other psycholinguistic variables and word recognition times. *Behavior Research Methods, 50,* 2366-2387.

Bonin, P., Méot, A., Aubert, L., Malardier, N., Niedenthal, P., & Capelle-Toczek, M.-C. (2003a). Normes de concrétude, de valeur

d'imagerie, de fréquence subjective et de valence émotionnelle pour 866 mots [Concreteness, imageability, subjective frequency and emotional valence norms for 866 words]. *L'Année Psychologique, 104,* 655-964.

Bonin, P., Méot, A., Boucheix, J.-M., & Bugaiska, A. (2018b). Psycholinguistic norms for 320 fixed expressions (idioms and proverbs) in French. *Quarterly Journal of Experimental Psychology, 71,* 1057-1069.

Bonin, P., Méot, A., Ferrand, L., & Bugaiska, A. (2015). Sensory experience ratings (SERs) for 1,659 French words: Relationships with other psycholinguistic variables and visual word recognition. *Behavior Research Methods, 47,* 813-825.

Bonin, P., Méot, A., Ferrand, L., & Roux, S. (2011). L'imageabilité : normes et relation avec d'autres variables psycholinguistiques. *L'Année Psychologique, 111,* 329-359.

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

Bronk, M., Zwitserlood, P., & Bölte, J. (2013). Manipulations of word frequency reveal differences in the processing of morphologically complex and simple words in German. *Frontiers in Psychology, 4:*546. **(Table, appendix)**

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,* 1129-1139.

Brysbaert, M., Mandera, P., McCormick, S. F., & Keuleers, E. (2019). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods, 51,* 467-479.

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science, 27,* 45-50.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.). *Language production: Development, writing and other language processes* (Vol. II, pp. 257–294). Academic Press.

Calting, J. C., & Elsherif, M. M. (2020). The hunt for the age of acquisition effect: It's in the links! *Acta Psychologica, 209,* 103138.

Chedid, G., Wilson, M. A., Bedetti, C., Rey, A. E., Vallet, G. T., & Brambati, S. M. (2019). Norms of conceptual familiarity for 3,596 French nouns and their contribution in lexical decision. *Behavior Research Methods, 51,* 2238-2247.

Chen, T.-M., & Chen, J.-Y. (2006). Morphological encoding in the production of compound words in Mandarin Chinese. *Journal of Memory and Language, 54,* 491–514. **(Table, appendix)**

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.

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,* 1072-1082.

Cortese, M. J., Yates, M., Schock, J., & Vilks, L. (2018). Examining word processing via a megastudy of conditional reading aloud. *Quarterly Journal of Experimental Psychology, 71,* 2295-2313.

Crepaldi, D., Rastle, K., Davis, C. J., & Lupker, S. J. (2013). Seeing stems everywhere: Position-independent identification of stem morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 510–525. **(Table, appendix)**

Cui, L., Häikiö, T., Zhang, W., Zheng, Y., & Hyönä, J. (2017). Reading monomorphemic and compound words in Chinese. *The Mental Lexicon, 12:1,* 1–20. **(Table, appendix)**

Cui, L., Yan, G., Bai, X., Hyönä, J., Wang, S., & Liversedge, S. P. (2013). Processing of compound-word characters in reading Chinese: An eye-movement-contingent display change study. *The Quarterly*

Journal of Experimental Psychology, 66, 527–547. **(Table, appendix)**

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65,* 171-237.

Dasgupta, T., Sinha, M., & Basu, A. (2016). Computational models of the representation of Bangla compound words in the mental lexicon. *Journal of Psycholinguistic Research, 45,* 833–855. **(Table, appendix)**

Davis, C. P., Libben, G., & Segalowitz, S. J. (2019). Compounding matters: Event-related potential evidence for early semantic access to compound words. *Cognition, 184,* 44–52. **(Table, appendix)**

Desrochers, A., & Thompson, G. L. (2009). Subjective frequency and imageability ratings for 3,600 French nouns. *Behavior Research Methods, 41,* 546-557.

Desrochers, A., Liceras, J. M., Fernández-Fuertes, R., & Thompson, G. L. (2010). Subjective frequency norms for 330 Spanish simple and compound words. *Behavior Research Methods, 42,* 109-117. **(Table, appendix)**

Dressler, W. U. (2006). Compound types. In G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 23–44). : Oxford University Press.

Drieghe, D., Pollatsek, A., Juhasz, B. J., & Rayner, K. (2010). Parafoveal processing during reading is reduced across a morphological boundary. *Cognition, 116,* 136–142. **(Table, appendix)**

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

Duñabeitia, J. A., Marín, A., Avilés, A., Perea, M., & Carreiras, M. (2009b). Constituent priming effects: Evidence for preserved morphological processing in healthy old readers. *European Journal of Cognitive Psychology, 21 (2/3),* 283-302. **(Table, appendix)**

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2007). The role of the frequency of constituents in compound words: Evidence from Basque and Spanish. *Psychonomic Bulletin & Review, 14,* 1171-1176. **(Table, appendix)**

Elsherif, M. M., Catling, J. C., & Frisson, S. (2020). Two words as one: A multi-naming investigation of the age-of-acquisition effect in compound word processing. *Memory & Cognition, 48,* 511-525. **(Table, appendix)**

Falkauskas, K., & Kuperman, V. (2015). When experience meets language statistics: Individual variability in processing English compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1607–1627. **(Table, appendix)**

Fang, S.-P. (2003). Are there differential word length effects in the two visual fields? *Brain and Language, 85,* 467-485. **(Table, appendix)**

Fehringer, C. (2012). The lexical representation of compound words in English: evidence from aphasia. *Language Sciences, 34,* 65–75. **(Table, appendix)**

Ferrand, L., Bonin, P., Méot, A., Augustinova, M., New, B., Pallier, C., & Brysbaert, M. (2008). Age-of-acquisition and subjective frequency estimates for all generally known monosyllabic French words and their relation with other psycholinguistic variables. *Behavior Research Methods, 40,* 1049–1054.

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

Ferrand, L., Méot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., Dufau, S., Mathôt, S., & Grainger, J. (2018). MEGALEX: A megastudy of visual and auditory word recognition. *Behavior Research Methods, 50,* 1285-1307.

Fiorentino, R., Naito-Billen, Y., Bost, J., & Fund-Reznicek, E. (2014). Electrophysiological evidence for the morpheme-based

combinatoric processing of English compounds. *Cognitive Neuropsychology, 31,* 123–146. **(Table, appendix)**

Fiorentino, R., & Poeppel, D. (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes, 22,* 953-1000. **(Table, appendix)**

Forgács, B., Bohrn, I., Baudewig, J., Hofmann, M. J., Pléh, C., & Jacobs, A. M. (2012). Neural correlates of combinatorial semantic processing of literal and figurative noun compound words. *NeuroImage, 63,* 1432–1442. **(Table, appendix)**

Frisson, S., Niswander-Klement, E., & Pollatsek, A. (2008). The role of semantic transparency in the processing of English compound words. *British Journal of Psychology, 99,* 87–107. **(Table, appendix)**

Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language, 60,* 20–35. **(Table, appendix)**

Gagné, C. L., & Spalding, T. L. (2016a). Effects of morphology and semantic transparency on typing latencies in English compound and pseudocompound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 1489-1495. **(Table, appendix)**

Gagné, C. L., & Spalding, T. L. (2016b). Written production of English compounds: Effects of morphology and semantic transparency. *Morphology, 26,* 133-135. **(Table, appendix)**

Gagné, C. L., Spalding, T. L., & Schmidtke, D. (2019). LADEC: The large database of English compounds. *Behavior Research Methods, 51,* 2152–2179. **(Table, appendix)**

Ghyselinck, M., Custers, R., & Brysbaert, M. (2004). The effect of age of acquisition in visual word processing: Further evidence for the semantic hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 550–554.

Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes, 15,* 421-444.

Gonthier, I., Desrochers, A., Thompson, G. L., & Landry, D. (2009). Normes d'imagerie et de fréquence subjective pour 1,760 mots monosyllabiques de la langue française. *Canadian Journal of Experimental Psychology, 63,* 139-149.

Güemes, M., Gattei, C., & Wainselboim, A. (2019). Processing verb-noun compound words in Spanish: Evidence from event-related potentials. *Cognitive Neuropsychology, 36,* 265-281. **(Table, appendix)**

Günther, F., & Marelli, M. (2019). Enter sandman: Compound processing and semantic transparency in a compositional perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45,* 1872-1882. **(Table, appendix)**

Günther, F., Marelli, M., & Bölte, J. (2020). Semantic transparency effects in German compounds: A large dataset and multiple-task investigation. *Behavior Research Methods, 52,* 1208-1224. **(Table, appendix)**

Häikiö, T., Bertram, R., & Hyönä, J. (2010). Development of parafoveal processing within and across words in reading: Evidence from the boundary paradigm. *The Quarterly Journal Of Experimental Psychology, 63,* 1982–1998. **(Table, appendix)**

Hasenäcker, J., & Schroeder, S. (2019). Compound reading in German: effects of constituent frequency and whole-word frequency in children and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45,* 920-933. **(Table, appendix)**

Huang, H.-W., Lee, C.-Y., Tsai, J.-L., & Tzeng, O. J.-L. (2011). Sublexical ambiguity effect in reading Chinese disyllabic compounds. *Brain and Language, 117,* 77–87. **(Table, appendix)**

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology, 16,* 104-127. **(Table, appendix)**

Hyönä, J., & Bertram, R. (2011). Optimal viewing position effects in reading Finnish. *Vision Research, 51,* 1279–1287. **(Table, appendix)**

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye- movement–contingent display change study. *Memory & Cognition, 32,* 523-532. **(Table, appendix)**

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1612-1627. **(Table, appendix)**

Inhoff, A. W., Bruhl, D., & Schwartz, J. (1996). Compound word effects differ in reading, on-line naming, and delayed naming tasks. *Memory & Cognition, 24,* 466-476. **(Table, appendix)**

Inhoff, A. W., Starr, M. S., Solomon, M., & Placke, L. (2008). Eye movements during the reading of compound words and the influence of lexeme meaning. *Memory & Cognition, 36,* 675-687. **(Table, appendix)**

Janssen, N., Bi, Y., & Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access for Mandarin Chinese and English compounds. *Language and Cognitive Processes, 23,* 1191–1223. **(Table, appendix)**

Janssen, N., Pajitas, P. E., & Caramazza, A. (2011). A set of 150 pictures with morphologically complex English compound names: Norms for name agreement, familiarity, image agreement, and visual complexity. *Behavior Research Methods 43,* 478–490. **(Table, appendix)**

Janssen, N., Pajitas, P. E., & Caramazza, A. (2014). Task influences on the production and comprehension of compound words. *Memory & Cognition, 42,* 780–793. **(Table, appendix)**

Jarema, G., Busson, C., Nikolova, R., Tsapkini, K., & Libben, G. (1999). Processing compounds: A cross-linguistic study. *Brain and Language, 68,* 362-369. **(Table, appendix)**

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition, 13,* 789-845.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131,* 684–712.

Juhasz, B. J. (2008). The processing of compound words in English: Effects of word length on eye movements during reading. *Language and Cognitive Processes, 23,* 1057–1088. **(Table, appendix)**

Juhasz, B. J. (2012). Sentence context modifies compound word recognition: Evidence from eye movements. *Journal of Cognitive Psychology, 24,* 855-870. **(Table, appendix)**

Juhasz, B. J. (2018). Experience with compound words influences their processing: An eye movement investigation with English compound words. *The Quarterly Journal of Experimental Psychology, 71,* 103-112. **(Table, appendix)**

Juhasz, B. J., & Berkowitz, R. N. (2011). Effects of morphological families on English compound word recognition: A multitask investigation. *Language and Cognitive Processes, 26,* 653-682. **(Table, appendix)**

Juhasz, B. J., Lai, Y. H., & Woodcock, M. L. (2015). A database of 629 English compound words: Ratings of familiarity, lexeme meaning dominance, semantic transparency, age of acquisition, imageability, and sensory experience. *Behavior research methods, 47,* 1004-1019. **(Table, appendix)**

Juhasz, B. J., Pollatsek, A., Hyönä, J., Drieghe, D., & Rayner, K. (2009). Parafoveal processing within and between words. *The Quarterly Journal of Experimental Psychology, 62,* 1356–1376. **(Table, appendix)**

Juhasz, B., Starr, M., Inhoff, A., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology, 94,* 223–244. **(Table, appendix)**

Juhasz, B. J., & Yap, M. J. (2013). Sensory experience ratings for over 5,
000 mono-and disyllabic words. *Behavior Research Methods, 45,*
160-168.

Juhasz, B. J., Yap, M. J., Dicke, J., Taylor, S. C., & Gullick, M. M.
(2011). Tangible words are recognized faster: The grounding of
meaning in sensory and perceptual systems. *The Quarterly
Journal of Experimental Psychology, 64,* 1683–1691.

Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye
movements and the use of parafoveal word length information in
reading. *Journal of Experimental Psychology: Human Perception
and Performance, 34,* 1560–1579. **(Table, appendix)**

Kaakinen, J. K., & Hyönä, J. (2010). Task effects on eye movements
during reading. *Journal of Experimental Psychology: Learning,
Memory, and Cognition, 36,* 1561–1566. **(Table, appendix)**

Kim, S. Y., Yap, M. J., & Goh, W. D. (2019). The role of semantic
transparency in visual word recognition of compound words: A
megastudy approach. *Behavior Research Methods, 51,* 2722–
2732. **(Table, appendix)**

Ko, I. Y., Wang, M., & Kim, S. Y. (2011). Bilingual reading of com-
pound words. *Journal of Psycholinguistic Research, 40,* 49–
73. **(Table, appendix)**

Koester, D., Gunter, Th. C., & Wagner, S. (2007). The morphosyntactic
decomposition and semantic composition of German compound
words investigated by ERPs. *Brain and Language, 102,* 64–
79. **(Table, appendix)**

Koester, D., Gunter, Th. C., Wagner, S., & Friederici, A. D. (2004).
Morphosyntax, prosody, and linking elements: The auditory pro-
cessing of German nominal compounds. *Journal of Cognitive
Neuroscience, 16:9,* 1647–1668. **(Table, appendix)**

Koester, D., & Schiller, N. O. (2008). Morphological priming in overt
language production: Electrophysiological evidence from Dutch.
NeuroImage, 42, 1622–1630. **(Table, appendix)**

Koester, D., & Schiller, N. O. (2011). The functional neuroanatomy of
morphology in language production. *NeuroImage, 55,* 732–
741. **(Table, appendix)**

Kuperman, V. (2013). Accentuate the positive: Semantic access in
English compounds. *Frontiers in Psychology, 4:203,* 1–
10. **(Table, appendix)**

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. (2009). Reading
polymorphemic dutch compounds: Toward a multiple route model
of lexical processing. *Journal of Experimental Psychology: Human
Perception and Performance, 35,* 876–895.

Kush, D., Dillon, B., Eik, R., & Staub, A. (2019). Processing of
Norwegian complex verbs: Evidence for early decomposition.
Memory & Cognition, 47, 335-350. **(Table, appendix)**

Li, M., Jiang, N., & Gor, K. (2017). L1 and L2 processing of compound
words: Evidence from masked priming experiments in English
Bilingualism: Language and Cognition, 20, 384–402. **(Table,
appendix)**

Liang, F., Blythe, H. I., Bai, X., Yan, G., Li, X., Zang, C., & Liversedge,
S. P. (2017). The role of character positional frequency on Chinese
word learning during natural reading. *PLoS ONE 12(11):*
e0187656. **(Table, appendix)**

Libben, G. (2014). The nature of compounds: A psychocentric perspec-
tive. *Cognitive Neuropsychology, 31,* 8–25.

Lim, R. Y. H., Yap, M. J., & Tse, C.-S. (2020). Individual differences in
Cantonese Chinese word recognition: Insights from the Chinese
Lexicon Project. *The Quarterly Journal of Experimental
Psychology, 73,* 504-518. **(Table, appendix)**

Lorenz, A., & Zwitserlood, P. (2016). Semantically transparent and
opaque compounds in German noun-phrase production: Evidence
for morphemes in speaking. *Frontiers in Psychology, 7:*1943.
(Table, appendix)

Lüttmann, H., Bölte, J., Böhl, A., & Zwitserlood, P. (2011). Evidence for
morphological composition at the form level of speech production.
Journal of Cognitive Psychology, 23, 818-836. **(Table, appendix)**

MacGregor, L. J., & Shtyrov, Y. (2013). Multiple routes for compound
word processing in the brain: Evidence from EEG. *Brain &
Language, 126,* 217–229. **(Table, appendix)**

Mankin, J. L., Thompson, C., Branigan, H. P., & Simner, J. (2016).
Processing compound words: Evidence from synaesthesia.
Cognition, 150, 1-9. **(Table, appendix)**

Marelli, M., & Luzzatti, C. (2012). Frequency effects in the processing of
Italian nominal compounds: Modulation of headedness and seman-
tic transparency. *Journal of Memory and Language, 66,* 644–
664. **(Table, appendix)**

Marelli, M., Zonca, G., Contardi, A., & Luzzatti, C. (2014). The repre-
sentation of compound headedness in the mental lexicon: A picture
naming study in aphasia. *Cognitive Neuropsychology, 31,* 26–39.
(Table, appendix)

Mok, L. W. (2009). Word-superiority effect as a function of semantic
transparency of Chinese bimorphemic compound words.
Language and Cognitive Processes, 24, 1039-1081. **(Table,
appendix)**

Muller, C. (1977). *Principes et méthodes de statistique lexicale.* Hachette.

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2014). Do masked
orthographic neighbor primes facilitate or inhibit the processing of
Kanji compound words? *Journal of Experimental Psychology:
Human Perception and Performance, 40,* 813–840. **(Table,
appendix)**

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A
new French lexical database. *Behavior Research Methods,
Instruments, & Computers, 36,* 516-524.

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.

Nicoladis, E. (2001). When is a preposition a linking element? Bilingual
children's acquisition of French compound nouns. *Folia
Linguistica, 36,* 45–63.

Nicoladis, E., & Krott, A. (2007). Word family size and french-speaking
children's segmentation of existing compounds. *Language
Learning, 57:2,* 201–228. **(Table, appendix)**

Ogawa, T., & Saito, H. (2006). Semantic activation in visual recognition
of Japanese two-kanji compound words: Interference and facilito-
ry effects of neighbors. *Psychologia, 49,* 162–177. **(Table,
appendix)**

Perret, C., & Bonin, P. (2019). Which variables should be controlled for
to investigate picture naming in adults? A Bayesian meta-analysis.
Behavior Research Methods, 51, 2533-2545.

Pham, H., & Baayen, H. (2015). Vietnamese compounds show an anti-
frequency effect in visual lexical decision. *Language, Cognition and
Neuroscience, 30,* 1077-1095. **(Table, appendix)**

Pollatsek, A., Bertram, R., & Hyönä, J. (2011). Processing novel and
lexicalised Finnish compound words. *Journal of Cognitive
Psychology, 23,* 795-810. **(Table, appendix)**

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in
the processing of Finnish compound words. *Language and
Cognitive Processes, 20,* 261–290. **(Table, appendix)**

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological
constituents in reading Finnish compound words. *Journal of
Experimental Psychology: Human Perception and Performance,
26,* 820-833. **(Table, appendix)**

Pollatsek, A., Slattery, T. J., & Juhasz, B. (2008). The processing of novel
and lexicalised prefixed words in reading. *Language and Cognitive
Processes, 23 (7/8),* 1133-1158. **(Table, appendix)**

Pratarelli, M. E. (1995). Modulation of semantic processing using word
length and complexity: an ERP study. *International Journal of
Psychophysiology, 19,* 233-246. **(Table, appendix)**

Reinhart, S., Schunck, A., Schaadt, A. K., Adams, M., Simon, A., &
Kerkhoff, G. (2016). Assessing neglect dyslexia with compound
words. *Neuropsychology, 30,* 869–873. **(Table, appendix)**

Roelofs, A. (1996). "Morpheme frequency in speech production: testing WEAVER." In: G. Booij & J. van Marle (Eds.), *Yearbook of Morphology* (pp. 135-154). Dordrecht: Kluwer Academic Publishers. **(Table, appendix)**

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent does not occur. *Quarterly Journal of Experimental Psychology, 42A,* 529–567. **(Table, appendix)**

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

Schmidtke, D., Gagné, C. L., Kuperman, V., Spalding, T. L., & Tucker, B. V. (2018a). Conceptual relations compete during auditory and visual compound word recognition. *Language, Cognition and Neuroscience, 33,* 923–942. **(Table, appendix)**

Schmidtke, D., Gagné, C. L., Kuperman, V., & Spalding, T. L. (2018b). Language experience shapes relational knowledge of compound words. *Psychonomic Bulletin and Review, 25,* 1468-1487.**(Table, appendix)**

Schmidtke, D., & Kuperman, V. (2019). A paradox of apparent brainless behavior: The time-course of compound word recognition. *Cortex, 116,* 250-267. **(Table, appendix)**

Schmidtke, D., Van Dyke, J. A., & Kuperman, V. (2018c). Individual variability in the semantic processing of English compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44,* 421–439. **(Table, appendix)**

Schulte im Walde, S., & Borgwaldt, S. R. (2015). Association norms for German noun compounds and their constituents. *Behavior Research Methods, 47,* 1199-1221.

Semenza, C., Arcara, G., Facchini, S., Meneghello, F., Ferraro, M., Passarini, L., Pilosio, C., Vigato, G., & Mondini, S. (2011). Reading compounds in neglect dyslexia: The headedness effect. *Neuropsychologia, 49,* 3116–3120. **(Table, appendix)**

Semenza, C., & Luzzatti, C. (2014). Combining words in the brain: The processing of compound words. Introduction to the special issue. *Cognitive Neuropsychology, 31,* 1-7.

Shen, W., Li, X., & Pollatsek, A. (2018a). The processing of Chinese compound words with ambiguous morphemes in sentence context. Quarterly Journal of Experimental Psychology, *71,* 131–139. **(Table, appendix)**

Shen, W., Li, X., & Tong, X. (2018b). Time course of the second morpheme processing during spoken disyllabic compound word recognition in Chinese. *Journal of Speech, Language, and Hearing Research, 61,* 2796–2803. **(Table, appendix)**

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin, 86,* 420.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory, 6,* 174-215.

Sohn, H.-M. (1999). *The Korean language.* Cambridge: Cambridge University Press.

Stockall, L., & Marantz, A. (2006). A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon, 1,* 85–123.

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology, 53,* 347–359.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15,* 607–620. **(Table, appendix)**

Tamaoka, K. (2007). Rebounding activation caused by lexical homophony in the processing of Japanese two-kanji compound words. *Reading and Writing, 20,* 413–439. **(Table, appendix)**

Tamaoka, K., & Hatsuzuka, M. (1998). The effects of morphological semantics on the processing of Japanese two-kanji compound words. *Reading and Writing: An Interdisciplinary Journal, 10,* 293–322. **(Table, appendix)**

Tse, C.-S., & Yap, M. J. (2018). The role of lexical variables in the visual recognition of two-character Chinese compound words: A megastudy analysis. The Quarterly Journal of Experimental Psychology, 71, 2022–2038. **(Table, appendix)**

Tse, C.-S.,Yap, M. J., Chan, Y.-L., Sze, W. P., Shaoul, C., & Lin, D. (2017). The Chinese Lexicon Project: A megastudy of lexical decision performance for 25,000+ traditional Chinese two-character compound words. *Behavior Research Methods, 49,* 1503–1519. **(Table, appendix)**

Verdonschot, R. G., La Heij, W., Tamaoka, K., Kiyama, S., You, W.-P., & Schiller, N. O. (2013). The multiple pronunciations of Japanese kanji: A masked priming investigation. *The Quarterly Journal of Experimental Psychology, 66,*2023–2038. **(Table, appendix)**

Vergara-Martínez, M., Duñabeitia, J. A., Laka, I., & Carreiras, M. (2009). ERP correlates of inhibitory and facilitative effects of constituent frequency in compound word reading. *Brain Research, 1257,* 53–64. **(Table, appendix)**

Wang, W., Lu, A., He, D., Zhang, B., & Zhang, J. X. (2017). ERP evidence for Chinese compound word recognition: Does morpheme work all the time? *NeuroQuantology, 15,* 142-152. **(Table, appendix)**

Yap, M. J., & Balota, D. A. (2015). Visual word recognition. In A. Pollatsek & R. Treiman (Eds.), *Oxford handbook of reading* (pp. 26-43). Oxford University Press.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognitions: Insights from the English lexicon project. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 53–79.

Yap, M. J., & Pexman, P. M. (2016). Semantic richness effects in syntactic classification: The role of feedback. *Frontiers in Psychology, 7,* 1394.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and language, 47*, 1-29.

Zevin, J.D., & Seidenberg, M.S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory and Cognition, 32,* 31-38.

Zhang, J., Anderson, R. C., Wang, Q., Packard, J., Wu, X., Tang, S., & Ke, X. (2012). Insight into the structure of compound words among speakers of Chinese and English. *Applied Psycholinguistics, 33,* 753–779. **(Table, appendix)**

Zhou, X., & Marslen-Wilson, W. (1994). Words, morphemes and syllables in the Chinese mental lexicon. *Language and Cognitive Processes, 9,* 393-422. **(Table, appendix)**

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes, 9,* 341–368. **(Table, appendix)**

Zwitserlood, P., Bölte, J., & Dohmes, P. (2000). Morphological effects on speech production: Evidence from picture naming. *Language and Cognitive Processes, 15,* 563–591.

Zwitserlood, P., Bölte, J., & Dohmes, P. (2002). Where and how morphologically complex words interplay with naming pictures. *Brain and Language, 81,* 358-367.