

Norms and comprehension times for 305 French idiomatic expressions

Patrick Bonin · Alain Méot · Aurelia Bugaiska

© Psychonomic Society, Inc. 2013

Abstract We report psycholinguistic norms for 305 French idiomatic expressions (Study 1). For each of the idiomatic expressions, the following variables are reported: knowledge, predictability, literality, compositionality, subjective and objective frequency, familiarity, age of acquisition (AoA), and length. In addition, we have collected comprehension times for each idiom (Study 2). The psycholinguistic relevance of the collected norms is explained, and different analyses (descriptive statistics, correlation and multiple regression analyses) performed on the norms are reported and discussed. The entire set of norms and reading times are provided as supplemental material.

Keywords Idiomatic expressions · Subjective frequency · Literality · Age of acquisition

A large number of psycholinguistic norms are available for individual pictures (of objects or of actions) and their names (for recent studies, see, e.g., Janssen, Pajtas, & Caramazza, 2011; Tsaparina, Bonin, & Méot, 2011). However, this is less frequently the case for longer linguistic items such as phrases, sentences, and idiomatic expressions. An idiomatic expression comprises several words whose overall meaning

(figurative meaning) is in some way different from the meanings of its constituent parts, or in other words, the meaning of a given expression is not a direct function of its components (Tabossi, Arduino, & Fanari, 2011). For instance the French idiomatic expression *casser sa pipe* (“to kick the bucket”) means “to die,” and thus its global meaning is far from the meaning of its components, which, taken separately, literally mean “to break (*casser*) his/her (*sa*) pipe (*pipe*).” In the present research, we pursued two main goals. The first was to collect norms for French idiomatic expressions in order to increase the pool of French norms available for this type of item. Unlike in the case of other types of stimuli (pictures, isolated words), norms for idiomatic expressions in French are very rare. Although Caillies (2009) reported norms for 300 expressions (only 35 of which were also used in the present study), she did not collect online measures of the processing of idioms. The second goal (and the novel aspect of our approach) was therefore to collect comprehension times for these expressions and to gain an understanding of the relationships between the various dimensions of idioms and their online processing. The new set of norms should facilitate research on idiom processing and may lead to new insights into how people represent and process idioms.

Idiomatic expressions exist in all languages and are very common (Cacciari & Tabossi, 1988)—in particular, in colloquial language (Tabossi, Fanari, & Wolf, 2008). These verbal expressions are often used in everyday conversation without apparent effort (Tabossi et al., 2008) and are also quickly recognized because they are well-known expressions (Tabossi, Fanari, & Wolf, 2009). Idioms are preformatted utterances, a special kind of fixed expression (other types of fixed expressions are phrasal verbs, restricted collocations, and sayings and proverbs) that belong to the conventional repertoire of native speakers of a given language (Sprengrer, Levelt, & Kempen, 2006). In effect, such utterances are not created by the speakers themselves, since both the meaning and form of these utterances are standardized, often allowing for only

Electronic supplementary material The online version of this article (doi:10.3758/s13428-013-0331-4) contains supplementary material, which is available to authorized users.

P. Bonin (✉) · A. Bugaiska
LEAD-CNRS (UMR 5022), University of Bourgogne, Pôle AAFE
- Esplanade Erasme, BP 26513, 21065 Dijon Cedex, France
e-mail: Patrick.Bonin@u-bourgogne.fr

P. Bonin
Institut universitaire de France, Paris, France

A. Méot
University Blaise Pascal, LAPSCO-CNRS, Clermont-Ferrand,
France

minimal variation. As a result, they cannot be easily translated from one language to another, since different expressions are used to represent the same idea, such as the English “raining cats and dogs,” for the idea that it is raining a lot, and the corresponding expression in French *pleuvoir comme vache qui pisse*, which literally translates as “raining like a pissing cow.” When learning a second language, speakers often face difficulties in mastering these expressions properly (Tabossi et al., 2009).

In psycholinguistics, a number of studies have attempted to investigate how idiomatic expressions are processed in different populations, such as healthy adults, patients (e.g., Hillert, 2004; Papagno & Caporali, 2007), or children (e.g., Caillies & Le Sourn-Bissaoui, 2008; Levorato, Nesi, & Cacciari, 2004), as well as in different tasks, such as comprehension (e.g., Caillies & Butcher, 2007; Caillies & Declercq, 2011) and production (e.g., Conner et al., 2011; Sprenger et al., 2006). The neural correlates of idiom processing have also been investigated (e.g., Hillert & Buracas, 2009). However, as of yet, only a few studies have been specifically dedicated to the collection of norms for idiomatic expressions. Such norms are available in English (Cronk, Lima, & Schweigert, 1993; Libben & Titone, 2008; Titone & Connine, 1994b) and in Italian (Tabossi et al., 2011). As was mentioned above, as far as French is concerned, we are aware of only one study that has reported norms for 300 idiomatic expressions (Caillies, 2009).

Theoretical issues addressed in the processing of idiomatic expressions

The collection of norms for idioms has made it possible to address theoretical issues pertaining to how idiomatic expressions are represented in memory, to the comprehension of idiomatic expressions, and to their verbal production (e.g., Cutting & Bock, 1997; Konopka & Bock, 2009; Sprenger et al., 2006; Tabossi et al., 2009). However, to be concise, we shall limit ourselves to certain theoretical aspects of idiom processing.

There are different views on the issue of how idioms are stored in memory. According to some researchers, idioms are represented as lexical units (i.e., word-like representations) or “giant,” morphologically complex words (e.g., Swinney & Cutler, 1979), whereas for others, they correspond to configurations of words (Cacciari & Tabossi, 1988). According to the configuration view, the meaning of an idiom is encoded in the weights of the connections between the lexical nodes that make up the idiom. For instance, the meanings of the individual words of “kick the bucket” are each represented by separate lexical nodes, and the meaning of the idiom is represented by the connections between “kick,” “the,” and “bucket.” More recently, Cutting and Bock (1997) proposed that whole

idioms are stored in the lexicon as nodes. These nodes, referred to as *superlemmas* (Sprenger et al., 2006), are located between concepts and word forms. The superlemmas are connected to phrasal nodes in the syntactic portion of the speech production system (e.g., “kick the bucket” is linked with a verb phrase), as well as to the individual lemmas making up the idioms (e.g., the lemmas of “kick” and “bucket”). The superlemmas are connected to a lexical concept node that represents the meaning.

Turning to the online processing of idioms, different views have been proposed to account for their comprehension. According to a noncompositional view, idioms are processed as long words (e.g., Swinney & Cutler, 1979). Within this perspective, the processes underpinning idiom recognition are the same as those mobilized during word recognition (Fanari, Cacciari, & Tabossi, 2010). Swinney and Cutler proposed the *lexical representation model*, according to which both the literal and figurative meanings of idioms are computed. The computation of the two types of meaning is initiated when the first word of the idiom is encountered. A strong version of this view is that people simply bypass literal meaning and directly access the figurative interpretation of the idiomatic string (Gibbs, 1984). The compositional view of idiom processing holds that understanding idiomatic meaning requires access to the internal semantic structure of the idiom (e.g., Gibbs, Nayak, & Cutting, 1989). Finally, there is a *hybrid* view that takes account of features from both the noncompositional and compositional views (Cacciari & Tabossi, 1988; Titone & Connine, 1999). According to the latter view, idioms have unitary representations that can be directly retrieved when the idioms are familiar or predictable. However, idioms may also be compositionally analyzed during comprehension, especially when they are unfamiliar or unpredictable. One version of the hybrid view—the configuration hypothesis (Cacciari & Tabossi, 1988)—holds that idioms are word configurations that are subjected to a linguistic analysis until a key part is processed that prompts recognition of the idiomatic nature of the input string and, consequently, results in the activation of its figurative meaning. It is therefore clear that no empirical test of the different views of idiom processing is possible without first collecting norms for idioms such as familiarity (and subjective frequency) and compositionality (and literalness). The different types of norms that are collected for idiomatic expression are now described.

What types of norms are collected for idiomatic expressions?

Norms for words are particularly useful to researchers since they make it possible to control for important variables in experimental studies. Thanks to the collection of norms for

pictures and their names, we know that image agreement, age of acquisition (AoA), and name agreement are among the most important variables to be taken into account when designing picture-naming experiments or when running multiple regression analyses on picture-naming performance (Alario et al., 2004; Bonin, Chalard, Méot, & Fayol, 2002). From a theoretical point of view, the collection of psycholinguistic norms has helped to delineate the processes and the representations that are involved in different lexical processing tasks, such as object naming (Alario et al., 2004; Bonin et al., 2002), lexical decision (Ferrand et al., 2011; Keuleers, Lacey, Rastle, & Brysbaert, 2012), or word reading (Ferrand et al., 2011). As far as idiomatic expressions are concerned, the collection of norms is also helpful when designing experimental studies aimed at investigating the processes underpinning idiomatic comprehension (e.g., Caillies & Butcher, 2007) or production (e.g., Cutting & Bock, 1997; Sprenger et al., 2006). On the basis of the use of norms for idiomatic expressions, as briefly reviewed above, several studies have been conducted to test a number of hypotheses pertaining to the mechanisms and the representations involved in the comprehension (and in the production) of idiomatic expressions and the way these are represented in memory.

Among the norms that have been collected for idioms, and that are described in turn below, are the following: knowledge, familiarity, subjective and objective frequency, AoA, predictability, literality, compositionality, and length. Below, we first briefly discuss the variables that we took into account in our study. The details relating to the data collection for each norm are presented in the “**Method**” section of **Study 1**.

One way to assess *knowledge of idioms* is to ask participants to provide a definition. Tabossi et al. (2011) used this original procedure in Italian participants. They had to explain the meaning of each idiom they knew, but not if they were unfamiliar with an item. The participants had to provide their explanation by using a single word, a sentence, or a longer formulation. Age effects have been found in the production of idiomatic expressions with the result that younger adults (ages 18–30 years) have been found to produce significantly more correct idiom responses than older adults aged 60–85 (Conner et al., 2011). Moreover, the older adults in this study produce proportionately more partial responses than do the younger adults even when the target idioms was equally familiar. According to Conner et al., in the same way as for single words, older adults experience retrieval difficulty with idioms. A potential weakness of the procedure used by Tabossi et al. (2011) is that, since it is a somewhat difficult and time-consuming task, the possibility cannot be excluded that certain participants may have answered that they did not know the definition of an idiom to avoid having to produce a definition. However, in Tabossi

et al.’s (2011) study, this phenomenon did not occur, as the reliable correlation found between familiarity and knowledge (.49) suggests. One strength of this procedure is that it directly assesses actual knowledge of an idiom and does not, therefore, rely on the personal conviction of the raters that they know the meaning. In our study, we used the latter procedure to evaluate the knowledge of such expressions. The participants were required to evaluate, on an *X*-point scale, the degree to which they thought they knew the meaning of the expression and could explain it verbally.

Predictability corresponds to the probability of completing an incomplete idiomatic expression. It is measured by asking participants to read a sentence fragment and then complete the missing part by writing down the first word that comes to mind (generally, on a line at the end of each fragment). Strong predictability effects have been reported in word reading, with the result that the processing of an upcoming word is facilitated in a predictable context, as compared with a less predictable one (e.g., Rayner, Ashby, Pollatsek, & Reichle, 2004). Also, in conceptually driven naming, the names of pictures are produced faster when the pictures are presented in the context of highly predictable sentences (e.g., the naming of the picture of a bed preceded by the sentence “He was tired so he went to ___”), as compared with a situation involving low-predictability sentences (e.g., “You will see a picture of a ___”) (Bonin, Roux, Barry, & Canell, 2012; Ferreira & Pashler, 2002; Griffin & Bock, 1998). As far as idiom processing is concerned, a cross-modal priming experiment found that predictability causes the (figurative) meaning of highly predictable idiomatic expressions to be accessed more quickly than that of less predictable expressions in comprehension (Cacciari & Tabossi, 1988). However, in the case of idioms rated as unpredictable, Cacciari and Tabossi found evidence for the immediate activation of the literal meaning only and late activation of both the literal and the figurative meanings (see also Titone & Connine, 1994a).

Some idioms have a well-formed literal meaning, whereas others possess only a meaningful idiomatic interpretation. *Literality* is a subjective measure of such a distinction between idioms. Originally, literal meaning was assumed to be conventional, compositional, and relatively context independent. Indeed, most idiomatic expressions are conventional (Caillies & Declercq, 2011). Titone and Connine (1994a) have provided evidence that the literal meaning of the last word of an idiom is activated. This is true for all idiom types, except for highly predictable nonliteral items.

Compositionality is a variable that refers to the fact that the constituents of some idiomatic expressions “carry identifiable parts of the idiomatic meaning.” *Decomposable idioms* correspond to expressions that comprise phrases whose individual components contribute to the overall meaning. By contrast, idioms whose individual words do

not contribute to the overall meaning are *nondecomposable idioms*. Caillies and Le Sourn-Bissaoui (2008) have shown that whereas young children (5 years old) can quickly understand decomposable idiomatic expressions in context, it is not until the age of 7–8 that they are able to understand nondecomposable idioms. Furthermore, in adult sentence reading, Caillies and Butcher (2007) found that decomposable idioms were processed sooner than nondecomposable ones. The compositionality of idioms is an interesting property that has been used to test several hypotheses related to the verbal processing of idioms. According to Tabossi, Wolf, and Koterle (2009), the processes underpinning the comprehension of decomposable idioms are different from those involved in that of nondecomposable items. These authors considered that the former type of idiom involves the same processes of lexical retrieval and syntactic parsing as those involved in the comprehension of literal expressions, whereas the latter mobilize processes similar to those underlying single-word recognition (Gibbs, Nayak, Bolton, & Keppel, 1989).

As far as isolated words are concerned, word familiarity or lexical familiarity (Gernsbacher, 1984), also often referred to as *subjective frequency* (Balota, Pilotti, & Cortese, 2001), have been thought to be a better index of the frequency of encounter of the words than is objective word frequency, and especially in the case of objective low-frequency words (Gernsbacher, 1984). However, more recent studies (e.g., Brysbaert & Cortese, 2011) have made it clear that subjective frequency ratings are useful primarily when objective word frequency counts are not of good quality. According to some researchers, subjective frequency ratings are no longer needed when good objective word frequency norms are available (Brysbaert et al., 2011; Brysbaert & Cortese, 2011). One problem with idiomatic expressions is that measures of the objective frequency of idioms are not as straightforward to establish as those that have become available for isolated words. Word frequency is certainly one of, if not the, most important variables to have been considered by researchers investigating reading or speaking. For instance, all influential models of word reading (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 2004) are able to account for the finding that high-frequency words are processed faster and more accurately than low-frequency words in experimental tasks such as lexical decision and reading aloud. However, the influence of objective word frequency in idiom processing has received little attention (Cronk et al., 1993; Libben & Titone, 2008). For instance, in the French normative study of 300 idiomatic expressions, this variable was not taken into account. As we shall explain below, there are different ways to take account of objective frequency measures when determining idiom frequency. If it is assumed that idiomatic meaning is built directly from its component

words and word meanings (i.e., the compositional view), any factor that has an influence on lexical word access and semantic retrieval should, in turn, facilitate composition of the idiomatic meaning. Thus, the frequency of the isolated words that make up the idiom should have an influence on idiom processing. Libben and Titone considered the verb word frequency and noun word frequency of idioms and found that they played different roles. In the present study, we took account of several objective frequency measures to assess the frequency of idiomatic expressions (namely, verb frequency and noun frequency, in line with Libben and Titone's work) and *cumulative* word frequency (namely, the summed frequencies of the individual words divided by the number of words, as used by Cronk et al.).

In normative studies of idiomatic expressions, the instructions for collecting familiarity norms sometimes relate specifically to the participant's degree of familiarity with each idiomatic expression, independently of whether or not he/she knows its meaning (Tabossi et al., 2011), or relate to the subjective frequency with which comprehenders encounter an idiom in its written or spoken form, regardless of their familiarity with the actual meaning of the phrase (Libben & Titone, 2008). Although these measures are certainly closely related, in our study, we thought it appropriate, following Titone and Connine (1994b), to distinguish between *the participant's own knowledge of the meaning* of an idiom and the *subjective frequency* of the idioms, defined as the frequency with which the participants think that they have encountered an idiom in its written or spoken form (also independently of their personal knowledge of its meaning). Familiarity has been found to have an impact in the reading of sentences containing idioms, with the result that sentences containing highly familiar idioms are processed faster than those involving low-familiarity idioms (Schweigert, 1986). Cronk and Schweigert (1992) found that high-literalness (idiom's literal meaning) idioms were read faster than low-literalness items, and the same effect was found when the familiarity of the figurative meaning of the idioms was taken into account. Also, Cronk et al. (1993) reported independent correlations of subjective frequency measures of idioms (one measure assessing how often [on a 1–5 scale] raters had heard the idioms used figuratively and another measure assessing how often [1–5 scale] they had heard them used literally). Furthermore, the familiarity of idioms is correlated with the ability to define them (Chan & Marinellie, 2008).

As far as the AoA of idiomatic expressions is concerned, this type of measure has not been taken into account in normative studies, apart from the Tabossi et al. (2011) study. By contrast, AoA measures have been frequently collected in normative studies of pictures and their names. The AoA of a word corresponds to the age at which a word is acquired in its written or spoken form, and the corresponding

measures are easily obtained by asking adults to provide an estimation on an X -point scale or in the form of a numerical value corresponding to an AoA. This variable is an important determinant of performance in various lexical processing tasks in different languages and populations (for reviews, see Johnston & Barry, 2006; Juhasz, 2005). It has repeatedly been found that words estimated to be learned early in life are processed faster and more accurately than words acquired later. AoA effects are stronger in tasks that obligatorily require semantic code activation, such as object or face naming, than in tasks that rely less on semantic codes, such as reading aloud (for comprehensive reviews, see Bonin, Barry, Méot, & Chalard, 2004; Johnston & Barry, 2006). Although the use of rated AoA to investigate age-limited learning effects in lexical processing has been criticized (Bonin et al., 2004; Bonin, Méot, Mermillod, Ferrand, & Barry, 2009; Zevin & Seidenberg, 2002, 2004), AoA norms are still being collected for large lists of words (e.g., Ferrand et al., 2008). Moreover, this variable is still used as a factor to investigate current issues in picture naming (e.g., De Zubicaray, Miozzo, Johnson, Schiller, & McMahan, 2012) or word recognition (e.g., Wilson, Ellis, & Burani, 2012). The collection of AoA ratings (as well as subjective frequency ratings) for idiomatic expressions might be thought to be less important for developmental studies of idiom processing, since it has been shown that the AoA of idioms is, to a large extent, dependent on the communication context (e.g., Cain, Oakhill, & Lemmon, 2005; Cain, Towse, & Knight, 2009). However, an interesting avenue of research for future studies could be to assess how well adult ratings of the AoA of idioms match the performance of children of different ages in idiom comprehension or production tasks. Since Tabossi et al. (2011) did not report reaction time (RT) data for idioms, it was not possible to study the influence of rated AoA in idiom comprehension in adults. However, this is possible in the present study.

Finally, the length variable is also taken into account in idiomatic norms. For isolated words, word length has been operationally defined in different ways: number of letters, phonemes, or syllables. Length effects have been reported in a few picture-naming studies (Alario et al., 2004) and in lexical decision studies (New, Ferrand, Pallier, & Brysbaert, 2006). The length of idiomatic expressions can be defined in terms of the number of words, syllables, or even letters or phonemes. In the normative study conducted by Tabossi et al. (2011), the length of idioms was defined in terms of the number of content words after the verb. Length effects in the processing of idioms have been reported. In a cross-modal study, Fanari et al., (2010) distinguished between short idioms (idioms with only one content word after the verb) and long idioms (more than one content word after the verb). In spoken comprehension, Fanari et al. used a cross-modal

visual lexical decision task and found that in the absence of a biasing context, the figurative meaning of an idiom was available at its offset for long but not for short strings. In the present study, we used two length measures—namely, the number of words and the number of letters.

Study 1: Collection of norms for idiomatic expressions

Method

Participants

A total of 187 participants (32 males; mean age, 20 years; range, 18–35 years) were involved. All were volunteers. Each participant took part in only one of the seven rating tasks. The different rating tasks were performed collectively. The participants had normal or corrected-to-normal vision.

Stimuli

The idiomatic expressions were selected from different Web sites dedicated to idiomatic expressions (e.g., <http://www.french-lessons.com/gallicismes2.html>). While we took care to choose expressions different from those used in the French study of Caillies (2009), there were 35 expressions in common. We selected both opaque and transparent idioms, as well as verbal and modal ones. When selecting the expressions, we retained those that we thought were most likely to be known by the majority of French undergraduate participants.

Procedure

Different booklets with all the idiomatic expressions were prepared. The participants were asked to rate the idioms on one of the following dimensions: knowledge, familiarity, subjective frequency, AoA, predictability, literality, or compositionality. The different instructions given to the participants in all the rating tasks are provided in full in [Appendix A](#) of the supplemental materials.

Knowledge ratings Each participant received a booklet containing the idiomatic expressions. For each idiomatic expression (e.g., *Marcher sur des œufs*—“To skate on thin ice”), the participants were instructed to rate, on a 5-point scale, the degree to which they knew the expression, with 1 = *absolutely no idea* and 5 = *absolutely certain of knowing its meaning and able to explain it verbally*.

Predictability scale The participants were told that they would read incomplete sentences and that their task would be to complete each sentence using one word only. For

example, they were shown the incomplete expression *Marcher sur des _____* (= “To skate on thin ice”) and had to fill in the blank with a word that made sense in the complete expression. For each idiom, the final rating measure was the proportion of participants who completed the sentence idiomatically.¹

Literality scale The participants were required to give a judgment concerning the literality of each of the idioms. All idioms have a figurative meaning, but only some of them also have a literal meaning. More precisely, for each idiomatic expression (e.g., *Marcher sur des œufs*, which for English readers resembles the idiom “to skate on thin ice” and translates literally as “to walk on eggs” and, in French, means to approach a situation or person with caution), the participants were instructed to rate on a 5-point scale, the degree to which they found that the expression had a plausible literal meaning, with 1 = *absolutely not plausible* and 5 = *completely plausible*. For instance, for the idiom *Marcher sur des œufs*, they may decide that it is not really plausible to “walk on eggs,” so the idiom has no literal sense. By contrast, the French expression *Jeter l’argent par les fenêtres* (= “To throw money down the drain”) may be judged to be completely plausible, because it is literally possible to throw money down the drain.

Compositionality scale Titone and Connine (1994b) used a sorting task to norm idioms on compositionality. Their participants had to sort the idioms into two categories, decomposable and nondecomposable idioms, and then further subdivide the decomposable idioms into two subcategories, *abnormally* and *normally* decomposable idioms. In

the present study, participants had to rate, on a 5-point scale, whether the idiom in each idiom-definition pair was decomposable (e.g., *avoir le pied marin* [have your sea legs] (= idiom) + *se dit d’une personne qui est à l’aise en mer et qui n’a pas le mal de mer* [said of a person who feels at home at sea and does not suffer from seasickness] (= definition)). The definitions of the idioms were taken from different sources—namely, from different Web sites on the Internet and from several dictionaries. They were additionally checked by the first and third authors for clarity of expression. A very small number of them were slightly rewritten to make them easier to understand.

They were told that an idiom is considered *decomposable* if its constituent parts contribute to the meaning of the expression. A rating of 1 was taken to indicate that the idiom was *not decomposable*, and 5 to indicate that the expression was *fully decomposable*. For example, the expression *Un éléphant dans un magasin de porcelaine* (= “a bull in a china shop” + the definition *désigne une personne qui intervient maladroitement dans une affaire délicate* [refers to someone who interferes clumsily in a delicate matter]) may be rated as 1 or 2 because the meanings of the different parts contribute little or nothing to the figurative meaning of the expression. By contrast, the expression *Jeter l’argent par les fenêtres* (= “to throw money down the drain” + the definition *dépenser de l’argent sans compter* [= to spend lavishly]) may be rated as 5 because the different parts contribute to the figurative meaning of the expression.

In addition to subjective measures, we took into account two types of objective measure: objective frequency and length of the idioms.

Subjective frequency ratings As far as the subjective frequency rating task is concerned, a 5-point scale was used to rate the frequency with which the participants thought they read, heard, or produced each idiom, with 1 = *never heard, read, or produced* and 5 = *heard, read, or produced very often*. The estimate had to be provided irrespective of the participants’ knowledge of the meaning of the expression.

Objective frequency measures Several objective frequency measures of idioms were computed. We took into account both verb frequency and noun frequency, as in Libben and Titone’s (2008) study. We also computed the summed frequencies of the individual words divided by the number of words, as in Cronk et al. (1993). For the latter measures, the inflected forms were taken into account. The frequency values were taken from the Lexique database (Lexique 3.71; New, Brysbaert, Véronis, & Pallier, 2007).

Familiarity For this scale, we adhered to the instructions provided by Tabossi et al. (2011). For each idiom (e.g., *Marcher sur des œufs* = “To skate on thin ice”), the

¹ Tabossi et al. (2011) added filler sentences to collect predictability norms, whereas we did not. The filler sentences were literal sentences created by replacing the idiomatic verb by another verb matched with it for frequency. Since we did not include literal sentences when collecting predictability scores, an anonymous reviewer questioned the validity of these scores. We therefore collected new predictability scores for our idiomatic expressions from an additional group of participants following the inclusion of literal sentences. For each idiomatic expression, a corresponding literal sentence was created on the basis of the procedure described by Tabossi et al. (2011). The entire list of idiomatic expressions was divided into two sublists, so that a given participant completed only half of the idiomatic expressions and a given idiom was never presented with its corresponding literal sentence. To anticipate the results, we found that (1) the correlation between the two norms was high, $r = .92$; (2) the means and the standard deviations were very similar ($M = .51$ and $.55$; $SD = .36$), (3) the only noticeable (but minor) differences concerned the correlations with orthographic length (.22 vs. .31) and RTs (.11 vs. .21); and importantly, (4) all the other analyses reported in the remainder of the article were strictly the same whatever the predictability scores used. In the remainder of the article, the reported analyses are those performed with the predictability scores obtained without literal sentences, since these were somewhat better correlated with word length. The two predictability scores are provided in Appendix B.

participants were instructed to judge how well they thought the expression was known by people like them, independently of whether or not they knew it themselves, with 1 = *the expression is unknown* and 5 = *the expression is known by everyone*.

Age of acquisition ratings For each idiomatic expression, the participants had to estimate the age at which they thought they had learned each expression (in its written or oral form). For this task, the five values of the scale corresponded to 3-year age bands with 0–3 at one extreme and 12+ at the other. The values were then converted to numerical values, with 1 = *learned between 0–3 years* and 5 = *learned at age 12 or after*.

Length variables In the normative study of Tabossi et al. (2011), the length of the idioms was defined in terms of number of content words after the verb. We computed this length measure for each idiom. However, in an attempt to better characterize the influence of length when idiomatic expressions are read, we also included several other length measures—namely, number of words and syllables and number of letters and phonemes. However, it turned out that the different length measures were highly correlated and that each length measure had comparable correlations with the other independent variables. For the sake of clarity, in the following, we therefore report only the analyses with number of letters taken as the length measure.

Results and discussion

The norms corresponding to each idiom are available in [Appendix B](#) of the supplemental material.

Reliability of collected norms

Table 1 reports the correlations between the different scores obtained from the means of 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].

As can be seen from Table 1, the reliability of the different norms was high. It can be noted, however, that the AoA variable was somewhat less reliable than the other norms.

Descriptive statistics

Table 2 shows that the means and the medians were all located in the right half of the scales. This was particularly the case for the knowledge variable, indicating that the participants thought that they knew the expressions, as well as the familiarity and subjective frequency variables.

Table 1 Correlation (r) between even and odd participants and ICC index

	N	$r(\text{even, odd})$	ICC index
Knowledge	30	.943	.964
Predictability	24	.932	
Literality	27	.933	.967
Compositionality	24	.905	.939
Subjective frequency	30	.932	.968
Familiarity	23	.927	.960
Age of acquisition	29	.811	.805

Note. Since for the predictability scale, the responses were binary, only the correlation between the percentages of correct answers for even and odd participants were computed. N = number of participants

The literality and decomposability distributions are nearly symmetrical, with more flat patterns as indicated by the interquartile range. As is illustrated by Fig. 1, a relatively pronounced negative skew was nevertheless observed for the knowledge, familiarity, and subjective frequency variables. This indicates that a few items were judged to be not well-known, familiar, or frequently encountered. It should be noted that the negative skew observed for the subjective frequency variable was not symmetrically observed on the objective frequency variables, for which, on the contrary, positive skews were observed. Finally, the AoA range is relatively small, as compared with the other variables, and possesses a positive skew.

Correlational analyses

As can be seen from Tables 3 (correlations) and 4 (factor analysis), two sets of variables can be distinguished. The first set corresponds to the knowledge, familiarity, and subjective frequency variables, which are highly and positively correlated. Compositionality and predictability are also strongly and positively correlated with these three first variables. In contrast, however, the correlation with AoA was high and negative, with the result that late-acquired idiomatic expressions were judged to be less well known and often less perceived as frequent, familiar, or predictable. Their individual components also contributed less to their overall meanings. A finding of more marginal significance was that the second component of the factor analysis was based primarily on the length of the expression, which was slightly positively correlated with the frequency of the noun appearing in the expression, AoA, and predictability, with the result that longer idioms tend to be acquired later in life and to be easier to complete. Finally, since objective *noun frequency* (log transformed) was, among other objective frequency

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

	Range	Mean	SD	Median	Q1	Q3	IRQ	Skewness
Knowledge	1–5	4.06	1.01	4.50	3.78	4.80	1.02	–1.43
Predictability	0–1	0.51	0.36	0.54	0.17	0.88	0.71	–0.10
Literality	1.04–4.93	2.68	1.16	2.54	1.62	3.59	1.97	0.33
Compositionality	1.21–4.87	2.93	0.89	2.88	2.21	3.67	1.46	0.11
Subjective frequency	1.27–4.90	3.58	0.94	3.83	3.07	4.30	1.23	–0.89
FF.sum	322–89,994	21,084	15,816	16,179	10,660	27,488	16,828	1.44
FF.mean	80–10,590	3,806	2,108	3,407	2,497	5,046	2,549	0.56
FF.noun	0–1,775	163	286	44	6	205	199	3.09
FF.verb	0–19,839	1,209	2,984	404	86	1,112	1,026	5.36
Familiarity	1.39–5	3.83	0.92	4.17	3.43	4.57	1.14	–1.05
AoA	1.71–5	2.98	0.66	2.80	2.47	3.33	0.86	0.94
Number of words	2–12	5.18	1.75	5	4	6	2	1.03
Number of letters	6–48	21.62	6.93	21	17	25	8	0.92

Note. AoA = age of acquisition; FF = film subtitle frequencies, values taken from Lexique 3.71 (New, Brysbaert, Véronis, & Pallier, 2007); FF.sum = sum of the FF of the words comprising the idiom; FF.mean = mean of the FF of the words; FF.noun = FF of the noun; FF.verb = FF of the verb. All frequency measures were log-transformed. SD = standard deviation; Q1 = 25th percentile; Q3 = 75th percentile; IQR = interquartile range

measures, the variable that was the most highly correlated with the other variables, we decided to report only information for objective noun frequency.²

Comparisons with other normative studies on idioms (French, Italian, and American-English)

In French, the correlations reported by Caillies (2009) between knowledge, predictability, compositionality, and familiarity³ had the same signs but, importantly, were smaller than those that were found to be reliable in the present study and are reported in Table 3. Only literality had comparable, but tenuous, relationships with the other variables. Turning to the items common to the two studies, it should first be remembered that there were relatively few of these (35). Moreover, a closer look at the expressions considered to be common to the two studies revealed that three expressions were not strictly the same (e.g., *envoyer au diable* vs. *aller au diable* [“devil take you”]; *faire l’autruche* vs. *pratiquer la politique de l’autruche* [“bury one’s head in the sand”]; *avoir du nez* vs. *avoir fin nez* [“have good

taste”]), while the pronoun *quelqu’un* (meaning “someone”) was added in six others (e.g., *être de mèche avec quelqu’un* [“to be in cahoots with”]). As a result, the number of identical expressions was too small to permit reliable comparisons between the two studies. It is also worth mentioning that certain rating tasks used in the Caillies study were not the same as the one used in the present study (i.e., knowledge and familiarity). In effect, Caillies used a definition task to assess knowledge of the items and the familiarity of the idioms was evaluated on a 6 point-scale (where 6 = *probably very familiar to you* and 1 = *not very familiar*). The only tasks that were the same were the predictability and literality rating tasks. We nevertheless performed a number of comparisons on

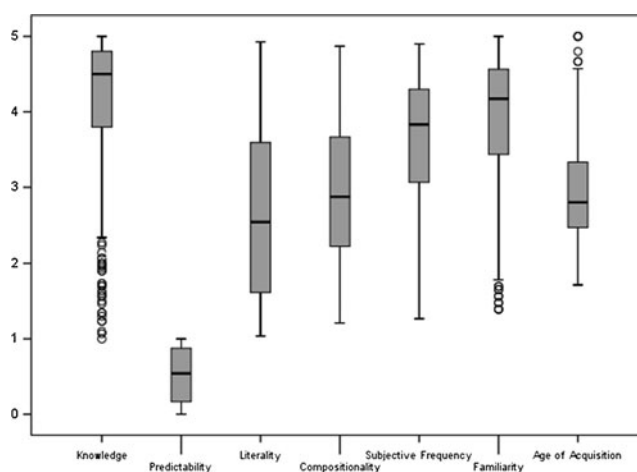


Fig. 1 Distributions of subjective ratings. Extreme values indicated by circles are between 1.5 and 3 interquartile ranges. Upper and lower horizontal segments correspond to the smallest and to the largest nonextreme values

² Indeed, we found that cumulative frequency was reliably correlated with length (.403, $p < .001$) and predictability (.168, $p < .05$), while mean frequency was significantly correlated with length (.175 $p < .05$). Finally, verb frequency was reliably correlated with predictability (–.158, $p < .05$).

³ It must be stressed that the definition of familiarity varies among researchers. Familiarity may correspond to what we have referred to as subjective frequency (e.g., Libben & Titone, 2008; Tabossi et al., 2011) or to a mixed measure of subjective frequency and conceptual familiarity (e.g., Caillies, 2009). Moreover, Libben and Titone (2008) and Titone and Connine (1994b) have also defined it in a second way—namely, as meaningfulness. Meaningfulness is roughly comparable to the knowledge measure used in the present study.

Table 3 Correlations between the measured variables

	2	3	4	5	6	7	8	9
1. Knowledge	.49***	.005	.59***	.86***	.21***	.86***	-.65***	-.054
2. Predictability		.08	.34***	.55***	.27***	.53***	-.37***	.22***
3. Literality				.15**	.07	.05	-.11	.04
4. Compositionality				.53***	.21***	.58***	-.43***	.09
5. Subjective frequency					.27***	.93***	-.74***	-.08
6. Objective noun frequency						.25***	-.15**	.19***
7. Familiarity							-.73***	-.05
8. Age of acquisition								.16**
9. Length (number of letters)								1.00

* $p < .05$ ** $p < .01$ *** $p < .001$

the similar expressions. These revealed that (1) for the predictability and literality tasks, the correlations were .70 and .91, respectively, and (2) the familiarity variable collected by Caillies correlated reliably with our familiarity (.63) and knowledge variables (.61).

As far as the Italian norms are concerned (Tabossi et al., 2011), we also found that the correlations were of the same sign and smaller than in the present norms with regard to knowledge, predictability, literality, compositionality, subjective frequency, AoA, and length. However, there were two noticeable exceptions. The first concerns the correlation between length and predictability, which was higher in Italian (.59) than in the present study. This is probably due to the different length measures that were used across the two studies (i.e., the number of words in Italian vs. the number of letters in the present study). The second is the correlation between literality and subjective frequency, which was negative in Italian (-.32) and not reliable in the present study.

In American English, Titone and Connine (1994b) reported norms for subjective frequency (first measure of familiarity, to use the authors' terminology), meaningfulness (a measure of familiarity roughly comparable to the knowledge variable in the present study), decomposability (an alternative measure of compositionality computed as the percentage of participants who judge the expression to be nondecomposable), predictability, and literality. The correlations of decomposability with all other variables and that between subjective frequency and predictability were noticeably lower in their study than in the present study. The correlations between the other pairs of variables were otherwise comparable.

If we now consider other correlations reported by Libben and Titone (2008) for another pool of idioms (using the same set of variables plus objective frequency), these were very similar to those obtained in the present study. The only difference worth mentioning concerns literality. This variable was correlated

negatively with meaningfulness in their study, whereas the corresponding correlation was not reliable in the present study. In addition, literality was negatively correlated with decomposability, whereas it was positively and reliably correlated in our data. However, the correlations were low (-.15 and -.21, respectively), and, as was mentioned above, the variables involved in the correlations were not exactly the same as those used here.

Study 2: Comprehension times for idiomatic expressions

Method

Participants

Sixty-one young adults taken from the same pool as in Study 1 took part in the online comprehension task.⁴ The comprehension task was performed individually. The participants were all volunteers and had normal or corrected-to-normal vision.

Stimuli

The idiomatic expressions normed in Study 1 were used.

Procedure

For the online idiom comprehension task, a Macintosh computer running the PsyScope software (Cohen, MacWhinney,

⁴ The reason why there were 61 participants in the idiom reading-comprehension task is because (1) half of the idioms were paired with a positive *yes* response and the remaining half with a negative *no* response and (2) any given participant was exposed to only half of the material with a given response type.

Flatt, & Provost, 1993) was used. The computer controlled the presentation of the expressions and their definitions and recorded the RTs. After a ready signal (a plus sign) presented for 300 ms in the center of the screen, an idiomatic expression was presented. The participants had to press the space bar with their nondominant hand as soon as they had read and understood the expression. Next, a definition was presented, and the participants had to decide as quickly as possible whether it matched the meaning of the idiom by pressing either a *yes* or a *no* response with one of the first two fingers of their dominant hand. The time taken to make the matched/nonmatched response was measured (in milliseconds). For half of the idioms, the idiomatic expression and the definition matched, and for the remaining half, they did not. Furthermore, we constructed two different lists, and any given participant saw only one of these. There were about 30 observations (RTs) for each idiom.

Results and discussion of Study 2

Scoring of the RT data

As far as online idiom comprehension is concerned, there were two RT measures for each expression: (1) the reading time and (2) the definition verification time. However, since the RTs obtained in the second (definition verification) task were extremely variable (e.g., 10 % below 1,159 ms and 10 % above 4,346 ms) and varied as a function of the definitions that were used, we do not report the corresponding analyses. Thus, only the RTs of the first (reading) online task corresponding to the trials on which the definition matched the idiomatic expression were retained in the analyses reported below. In the following, we refer to these measures as *reading times*. Thus, trials on which participants responded that the definition did not match the idiomatic expression were eliminated.

Table 4 Variable loadings on the first two components of a factor analysis (PCA)

	1	2
Knowledge	.90	
Predictability	.64	.37
Literality		
Compositionality	.69	
Subjective frequency	.94	
Objective noun frequency	.35	.55
Familiarity	.94	
Age of acquisition	-.79	.26
Length (number of letters)		.86

Note. The two components solution explains 61.7 % (47.2 % and 14.6 %) of the total variance. Since the third eigenvalue was 1.02, its associate component was not retained. Loadings less than .2 are not reported.

There were two expressions for which the definitions were found a posteriori to be unsuitable, and these were therefore excluded (i.e., *avoir une sale tête* and *être en cheville avec quelqu'un*). Twenty-two items for which more than 50 % of the participants indicated that the corresponding definitions did not match the meanings were also eliminated. Among the 8,571 RT trials recorded for the 281 remaining items, 801 (9.3 %) gave rise to incorrect decisions about the matching of the definition to the expression. Moreover, 22 trials (0.3 %) were affected by technical problems. RTs lower than 200 ms and higher than 3,500 ms were also eliminated (1.4 % and 1.4 % of the remaining RTs). Finally, RTs three standard deviations below or above the item means were excluded (0.8 % of the data). Overall, 12.8 % of the RTs were discarded from the 281 retained items.

The reading times corresponding to each idiom are available in [Appendix B](#) of the supplemental material.

The correlations with the knowledge, familiarity, and subjective frequency variables were negative (see [Table 5](#)), thus suggesting that familiar and more frequently encountered expressions are read faster than less known, less familiar, and less frequently encountered expressions. Finally, there was a slight but reliable positive correlation with the noun frequency of the idioms. This correlation, which was also observed when the sum or the mean of the frequencies of the words comprising the expression was taken into account, was also low. Given that the correlations involving objective frequency measures are less high than those involving the subjective frequency measure, this suggests that the objective frequency measures computed for idiomatic expressions are a less valid index of the true frequencies of encounter of the expressions. In the word reading literature, it is assumed that the higher the correlation of a given word frequency measure is with RT data (e.g., lexical decision), the better this measure indexes the true frequency with which the participants have encountered the words (Dimitropoulou, Duñabeitia, Avilés, Corral, & Carreiras, 2010). We follow the same line of reasoning here.

As is shown in [Table 5](#), the correlation between the number of letters in the expression and reading times was very high. Not surprisingly, and in line with previous findings (e.g., Fanari et al., 2010), long idiomatic expressions took longer to read than shorter ones. The relationships between reading times and the other variables were tenuous, with the most important being with AoA, with the result that early-acquired idioms were read more rapidly than late-acquired idioms.

A hierarchical regression analysis was conducted with reading times as the dependent variable. A first set of independent variables comprised subjective frequency, AoA, and length (number of letters). The first two of these variables were entered in the first set because they are commonly used when analyzing

Table 5 Correlations between comprehension times and the measured variables

Knowledge	Predictability	Literality	Compositionality	Subjective frequency	FF.noun	Familiarity	AoA	Length
-.233***	.111	-.049	-.004	-.260***	.205**	-.255***	.317***	.844***

Note. FF.noun = film subtitle frequency of the noun (log-transformed); AoA = age of acquisition ; Length = number of letters

* $p < .05$

** $p < .01$

*** $p < .001$

RT in word recognition studies (e.g., Brysbaert & Cortese, 2011; Morrison & Ellis, 1995). Subjective frequency was preferred over objective frequency because, as was pointed out above, the latter does not seem to be a good measure of the frequency of encounter of complete idiomatic expressions. As Table 6 indicates, only AoA and length were reliable predictors of reading times. The observation that AoA reliably predicts the time taken to read and comprehend idioms is a very important one, since AoAs have as yet only been infrequently collected for idioms. For example, Tabossi et al. (2011) collected AoA norms for idioms in Italian, but, since they did not collect online measures for these expressions, they were not able to assess the influence of rated AoA in idiom comprehension. The influence of length in the understanding of idioms is, however, not really surprising, and this finding has already been reported (e.g., Fanari et al., 2010).

Literality, familiarity and knowledge were entered in a second step using a stepwise procedure. This procedure was chosen as a result of (1) the uncertainty that exists about the reliability of the effects of these variables and (2) the high correlations that were found between some of them. Literality was the first variable to enter, followed by familiarity and knowledge. When reliable, these variables added a very low percentage of explained variance (between 0.4 % and 0.7 %). Concerning literality and familiarity, the findings are consistent with previous reports of an influence of

these factors (e.g., Schweigert, 1986, and Titone & Connine, 1994a, respectively). As far as familiarity is concerned, it should be remembered that its definition varies between researchers (see Footnote 3). In our study, familiarity corresponded to the participants' judgement of how well the expression was known by people like them (independently of their own knowledge). It is worth mentioning that we did not find a reliable contribution of predictability in the time taken to read and understand the idiom, in contrast to previous findings obtained in factorial experiments (e.g., Cacciari & Tabossi, 1988). There was also no reliable effect of compositionality. This contrasts with the finding of Caillies and Butcher (2007), who observed an effect of this variable in adults' sentence reading.

Libben and Titone (2008, Experiment 2) used a regression approach with the time taken to say whether an expression was meaningful as a dependent variable and subjective frequency, decomposability, the interaction between these two variables, literality, noun frequency, and verb frequency as independent variables. It is worth mentioning that in Libben and Titone's Experiment 2, length variables were not included in the regression model, because these were controlled for by using fixed length expressions. The only reliable variable in Libben and Titone's Experiment 2 was subjective frequency. The two objective frequency measures used were only marginally significant. Although the RT and compositionality measures used in our present study were the only variables that were roughly comparable to the variables used in Libben and Titone's study, a similar analysis conducted in French yielded reliable effects of subjective frequency, $b = -.46$, $t(274) = -6.18$, $p < .001$, compositionality, $b = .15$, $t(274) = 2.27$, $p < .05$, and objective noun frequency, $b = .28$, $t(274) = 4.78$, $p < .001$. This positive effect of objective noun frequency suggests that this type of objective measure is a poor index of the genuine frequency of encounter of idiomatic expressions.

Table 6 Regression analysis on comprehension times

		Betas	ΔR^2
Set 1 (simultaneous)	Subjective frequency	.01	.000
	Age of acquisition	.11*	.005
Set 2 (stepwise)	Letters	.83***	.654
	Knowledge	.14*	.004
	Literality	-.07*	.005
	Familiarity	-.23**	.007

Note. $\Delta R^2 = R$ -square change when the variable was entered as the last one. R^2 was .758 (.745 with only the variables of set 1)

* $p < .05$

** $p < .01$

*** $p < .001$

Conclusion

We have collected knowledge, predictability, literality, compositionality, familiarity, subjective frequency, and AoA

norms for a new set of French verbal idiomatic expressions (Study 1). We have also provided reading-comprehension times for these expressions (Study 2). The present work helps to increase the pool of available norms for French idioms. The psycholinguistic norms provided here are known to influence the processing of these expressions. Indeed, we found that the length and AoA of the idioms were the strongest predictors of adult reading times. These norms will be useful for the design of experimental studies aimed at testing a range of issues relating to the processing of idiomatic expressions (e.g., how these preformatted utterances are stored and processed in both comprehension and production tasks). Indeed, the processing of idiomatic expressions is an important area of research in psycholinguistics, since these types of expressions are very common in language (Cacciari & Tabossi, 1988) and are frequently used (Tabossi et al., 2008) and heard in everyday conversation.

Author Note This research was supported by a grant from the Institut universitaire de France to Patrick Bonin. We thank two anonymous reviewers and Prof. Gregory Francis for their very helpful comments on a previous version of the manuscript. We are very grateful to Sophie Pasquet and Mélanie Provost for their help in the preparation of the stimuli and the questionnaires and for running the comprehension experiment.

References

- Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*, 140–155.
- Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, *29*, 639–647.
- Bonin, P., Méot, A., Mermillod, M., Ferrand, L., & Barry, C. (2009). The effects of age of acquisition and frequency trajectory on object naming: Comments on Perez (2007). *The Quarterly Journal of Experimental Psychology*, *62*, 1132–1140.
- 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., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*, 89–114.
- Bonin, P., Roux, S., Barry, C., & Canell, L. (2012). Evidence for a limited-cascading account of written word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1741–1758.
- Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology*, *64*, 545–559.
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*, 412–424.
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, *27*, 668–683.
- Caillies, S. (2009). Descriptions de 300 expressions idiomatiques: Familiarité, connaissance de leur signification, plausibilité littérale, “décomposabilité” et “prédicibilité”. *L'Année Psychologique*, *109*, 463–508.
- Caillies, S., & Butcher, K. (2007). Comprehension of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol*, *22*, 79–108.
- Caillies, S., & Declercq, C. (2011). Kill the song—steal the show: What does distinguish predicative metaphors from decomposable idioms? *Journal of Psycholinguistic Research*, *40*, 205–223.
- Caillies, S., & Le Sourn-Bissaoui, S. (2008). Children’s understanding of idioms and theory of mind development. *Developmental Science*, *11*, 703–711.
- Cain, K., Oakhill, J., & Lemmon, K. (2005). The relation between children’s reading comprehension level and their comprehension of idioms. *Journal of Experimental Child Psychology*, *90*, 65–87.
- Cain, K., Towse, A. S., & Knight, R. S. (2009). The development of idiom comprehension: An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology*, *102*, 280–298.
- Chan, Y.-L., & Marinellie, S. A. (2008). Definitions of idioms in pre-adolescents, adolescents, and adults. *Journal of Psycholinguistic Research*, *37*, 1–20.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers*, *25*, 257–271.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.
- Conner, P. S., Hyun, J., O’Connor Wells, B., Anema, I., Goral, M., Monéreau-Merry, M.-M., ... Obler, L. K. (2011). Age-related differences in idiom production in adulthood. *Clinical Linguistics & Phonetics*, *25*, 899–912.
- Cronk, B. C., & Schweigert, W. A. (1992). The comprehension of idioms: The effects of familiarity, literalness, and usage. *Applied PsychoLinguistics*, *13*, 131–146.
- Cronk, B. C., Lima, S. D., & Schweigert, W. A. (1993). Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research*, *22*, 59–92.
- Cutting, J. C., & Bock, K. (1997). That’s the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition*, *25*, 57–71.
- De Zubicaray, G. I., Miozzo, M., Johnson, K., Schiller, N. O., & McMahan, K. L. (2012). Independent distractor frequency and age-of-acquisition effects in picture-word interference: fMRI evidence for post-lexical and lexical accounts according to distractor type. *Journal of Cognitive Neuroscience*, *24*, 482–495.
- Dimitropoulou, M., Duñabeitia, J. A., Avilés, A., Corral, J., & Carreiras, M. (2010). Subtitle-based word frequencies as the best estimate of reading behavior: The case of Greek. *Frontiers in Psychology*, *1*, 218.
- Fanari, R., Cacciari, C., & Tabossi, P. (2010). The role of idiom length and context in spoken idiom comprehension. *European Journal of Cognitive Psychology*, *22*, 321–334.
- Ferrand, L., Bonin, P., Méot, A., Augustinova, M., New, B., Pallier, C., et al. (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., Pallier, C. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in Psychology*, *2*:306.
- Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1187–1199.

- Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*, 256–281.
- Gibbs, R. W. (1984). Literal meaning and psychological theory. *Cognitive Science*, *8*, 275–304.
- Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989a). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language*, *28*, 576–593.
- Gibbs, R. W., Nayak, N. P., Bolton, J. L., & Keppel, M. E. (1989b). Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition*, *17*, 58–68.
- Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, *38*, 313–338.
- Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720.
- Hillert, D. G. (2004). Spared access to idiomatic and literal meanings: A single-case approach. *Brain and Language*, *89*, 207–215.
- Hillert, D. G., & Buracas, G. T. (2009). The neural substrates of spoken idiom comprehension. *Language & Cognitive Processes*, *24*, 1370–1391.
- Janssen, N., Pajtas, 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.
- Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition*, *13*, 789–845.
- Juhász, B. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, *131*, 684–712.
- Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*, 287–304.
- Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? structural generalisation from idiom production. *Cognitive Psychology*, *58*, 68–101.
- Levorato, M. C., Nesi, B., & Cacciari, C. (2004). Reading comprehension and understanding idiomatic expressions: A developmental study. *Brain and Language*, *91*, 303–314.
- Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, *36*, 1103–1121.
- Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 116–133.
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*, 661–677.
- New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Re-examining word length effects in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, *13*, 45–52.
- Papagno, C., & Caporali, A. (2007). Testing idiom comprehension in aphasic patients: The effects of task and idiom type. *Brain and Language*, *100*, 208–220.
- Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 720–732.
- Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research*, *15*, 33–45.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428.
- Sprenger, S. A., Levelt, W. J. M., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, *54*, 161–184.
- Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, *18*, 523–534.
- Tabossi, P., Arduino, L., & Fanari, R. (2011). Descriptive norms for 245 Italian idiomatic expressions. *Behavior Research Methods*, *43*, 110–123.
- Tabossi, P., Fanari, R., & Wolf, K. (2008). Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 313–327.
- Tabossi, P., Fanari, R., & Wolf, K. (2009a). Why are idioms recognized fast? *Memory & Cognition*, *37*, 529–540.
- Tabossi, P., Wolf, K., & Koterle, S. (2009b). Idiom syntax: Idiosyncratic or principled? *Journal of Memory and Language*, *61*, 77–96.
- Titone, D. A., & Connine, C. M. (1994a). Comprehension of idiomatic expressions: Effects of predictability and literalness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1126–1138.
- Titone, D. A., & Connine, C. M. (1994b). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literalness. *Metaphor and Symbolic Activity*, *9*, 247–270.
- Titone, D. A., & Connine, C. M. (1999). On the compositional and compositional nature of idiomatic expressions. *Journal of Pragmatics*, *31*, 1655–1674.
- Tsaparina, D., Bonin, P., & Méot, A. (2011). Russian norms for name agreement, image agreement for the colorized version of the Snodgrass and Vanderwart pictures and age of acquisition, conceptual familiarity, and imageability scores for modal object names. *Behavior Research Methods*, *43*, 1085–1099.
- Wilson, M. A., Ellis, A. W., & Burani, C. (2012). Age-of-acquisition affects word naming in Italian only when stress is irregular. *Acta Psychologica*, *139*, 417–424.
- 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.