Descriptive norms for 245 Italian idiomatic expressions

Patrizia Tabossi · Lisa Arduino · Rachele Fanari

Published online: 20 November 2010
© Psychonomic Society, Inc. 2010

Abstract The present study reports descriptive normative measures for 245 Italian verbal idiomatic expressions. For each of the idiomatic expressions the following variables are reported: Length, Knowledge, Familiarity, Age of Acquisition, Predictability, Syntactic flexibility, Literality and Compositionality. Syntactic flexibility was assessed using five syntactic operations: adverb insertion, adjective insertion, left dislocation, passive and movement. The psycholinguistic relevance of each dimension, their measures and the correlations among them are provided and discussed. The databases are freely available for down-loading from the Psychonomic Society Web archive at www.psychonomic.org/archive/.

Keywords Idiomatic expressions · Descriptive norms · Italian

Electronic supplementary material The online version of this article (doi:10.3758/s13428-010-0018-z) contains supplementary material, which is available to authorized users.

P. Tabossi
University of Trieste,
Trieste, Italy

L. Arduino
LUMSA University and ISTC-CNR,
Roma, Italy

R. Fanari
University of Cagliari,
Cagliari, Italy

P. Tabossi (✉)
Dipartimento di Psicologia,
Via S. Anastasio, 12,
34134 Trieste, Italy
e-mail: tabossi@units.it

Idiomatic expressions are usually characterized as linguistic expressions whose meaning is not a direct function of the meanings of their components: according to the syntax of English, the figurative meaning of *kick the bucket* (“to die”) cannot be derived by combining the meaning of *kick*, *the*, and *bucket*. As a consequence, the comprehension and production of these expressions are not adequately explained by standard models of language comprehension and production, which typically rely on the principle of semantic compositionality. In fact, they call for specific models, and much research has been conducted to elucidate how idioms are recognized and produced in their citation form (Bobrow & Bell, 1973; Cacciari & Tabossi, 1988; Gibbs, Nayak, & Cutting, 1989; Swinney & Cutler, 1979; Titone & Coninne, 1999).

Another characteristic of idioms is that they are syntactically defective, i.e. they do not accept all the syntactic operations normally allowed by literal strings with the same structure. For example, “*The bucket was kicked by John*” loses its figurative meaning for most speakers of English. How people deal with the syntax of idiomatic expressions both in comprehension and production is an open issue (Cutting & Bock, 1997; Gibbs & Gonzales, 1985; Gibbs & Nayak, 1989; Konopka & Bock, 2009; Sprenger, Levelt, & Kempen, 2006; Tabossi, Wolf, & Koterle, 2009).

A final important characteristic of idioms is that they are very common, particularly in spoken language. They constitute a relevant proportion of the vast family of formulaic expressions, which include names of people, places (e.g., *Beverly Hills*), clichés (e.g., *Love conquers all*), titles of songs, books (e.g., *All you need is love*), quotations (e.g., *Beam me up, Scotty*), foreign phrases (e.g., *c’est la vie*) (Jackendoff, 1995; Langacker, 1987). These expressions are estimated to be as numerous as individual words (Jackendoff, 1995) and frequently used. In particular

with respect to idioms, speakers show a strong propensity to “speak idiomatically unless there is a good reason not to do so” (Searle, 1975). In light of these observations, it is clear that any adequate theory of language comprehension and production must explain how people deal with idiomatic expressions.

Over the years, much progress has been made, and the view on idioms has changed. For a long time linguists and psycholinguists treated idiomatic expressions as long words, syntactically frozen and semantically non modifiable (Katz, 1973; Swinney & Cutler, 1979). To date, most researchers agree that idioms are not processed as lexical items and form a very heterogeneous class, which varies in terms of syntactic flexibility and semantic analyzability (Cacciari & Glucksberg, 1991; Cacciari & Tabossi, 1988; Gibbs, 1992; Glucksberg, 1993; Nunberg, Sag, & Wasow, 1994). Various models have also been proposed to explain how these expressions are mentally represented and processed (Cutting & Bock, 1997; Gibbs & Nayak, 1989; Sprenger et al., 2006; Tabossi, Fanari, & Wolf, 2009; Tabossi, Wolf, & Koterle, 2009). Also, interesting research is conducted on how children acquire the ability to use idiomatic expressions (Ackerman, 1982; Cacciari & Levorato, 1989; Gibbs, 1987, 1991; Levorato & Cacciari, 1995; Levorato, Nesi, & Cacciari, 2004), how idiom knowledge and processing are disturbed in aphasic patients (Hillert, 2004; Papagno, Curti, Rizzo, Crippa, & Colombo, 2006; Papagno, Tabossi, Colombo, & Zampetti, 2004), and what structures in the brain are responsible for the processing of idioms and other types of figurative language (Sidtis Van Lacker, 2006).

All this work has made clear that idiom processing is affected by a variety of variables. However, very little normative data is available (Cronk, Lima, & Schweigert, 1993; Titone & Connine, 1994b; Libben & Titone, 2008), and none in Italian. In this study, values referring to several variables for 245 Italian verbal idioms are reported. The variables have been selected on the basis of their psycholinguistic relevance and, loosely speaking, are organized from (more) superficial to (more) semantic ones. The present database, available on the Psychonomic Society Web archive, provides a useful instrument for research on idiom representation and processing in Italian. In what follows, we first discuss briefly the variables we have considered. Then, we present the database and the methods of data collection. Finally, we conduct a correlation analysis among the variables and discuss the results.

Length

Verbal idiomatic expressions may differ in length. Some strings contain only a content word after the verb (e.g., *alzare il gomito/raise the elbow*, “drink too much alcohol”),

others are much longer (e.g., *mettere una pulce nell'orecchio/put a flea in the ear*, “raise suspicion”). Several cross-modal studies have shown that the activation of the figurative meaning of an idiomatic string may occur rather late during the processing of an auditory presented string, and may not yet be available at its offset (Cacciari & Tabossi, 1988; Cacciari, Padovani, & Corradini, 2007; Tabossi & Zardon, 1993, 1995; Tabossi, Fanari, & Wolf, 2005). However, other data seem to contradict these results (Titone & Connine 1994a; Tabossi, Fanari, & Wolf, 2005). In a recent cross-modal study, verbal idioms were divided into short (i.e., containing only one content word after the verb) and long (e.g., containing more than one content word after the verb). The results indicated that in the absence of a biasing context, other things being equal, the figurative meaning of an idiom is available at its offset if the string is long, but it is not available if the string is short (Fanari, Cacciari, & Tabossi, 2010). Evidently, length may be a relevant factor to keep in mind when selecting idioms to use in processing studies.

Knowledge

The most complete norms on idioms published so far are Titone and Connine (1994b), recently upgraded and extended by Libben and Titone (2008). In both corpora, two measures of familiarity are included. The first measure refers to the subjective frequency of the expressions, and participants evaluate how frequently they encountered the idioms. The second measure requires participants to judge how meaningful each expression is. According to the authors, the second scale indicates the degree to which each idiom is understandable to participants. We agree with Titone and colleagues that both variables are relevant, and decided to consider them separately. Therefore, we collected a measure of knowledge of each expression independently from familiarity.

Familiarity

Familiarity has been defined as the frequency with which a listener or reader encounters a word in its spoken or written form (Gernsbacher, 1984). This variable is well-known to influence word recognition, and it is a strong predictor of speed and accuracy in several experimental tasks such as lexical decision and reading aloud (Connine, Mullennix, Shernoff, & Yelen, 1990; Ellis, 2002; Gernsbacher, 1984).

Familiarity has also been explored in idiom processing. For example, Schweigert (1986, 1991) showed that reading sentences is faster if they contain high familiarity rather than low familiarity idioms. Cronk and Schweigert (1992) investigated the relation between familiarity and literality

(see below for literality). They found that idiomatic expressions in their figurative interpretation are read faster when they have high rather than low familiarity.

Tabossi, Fanari, & Wolf (2009) explored the so-called *idiom superiority effect*: deciding that an idiom (e.g., *break the ice*) is a meaningful expression takes less time than deciding that a matched literal expression (e.g., *break the glass*) is meaningful. The phenomenon is well-known (Gibbs, 1980; McGlone, Glucksberg, & Cacciari, 1994; Ortony, Schallert, Reynolds, & Anton, 1978; Swinney, & Cutler, 1979), but its interpretation differs in the various models (Swinney & Cutler, 1979; Gibbs, Nayak, & Cutting, 1989; Cacciari & Tabossi, 1988). Tabossi et al. (2009) showed that idiomaticity *per se* plays no role in the effect, which is due to the familiarity of the expressions, whether idiomatic or not. These studies indicate the relevance of familiarity when investigating idiom processing.

Age of Acquisition (AoA)

Age of acquisition is defined as the estimated age at which a word and its meaning were first learned, either in the auditory or in the written modality. Typically, it is obtained by asking participants to evaluate at what age they think they have learnt the word (Morrison & Ellis, 1995, 2000). Although this methodology provides a subjective estimate of AoA, several studies have shown a high correlation between the ratings and objective, independent measures of age of acquisition suggesting that they provide reliable information (Lyons, Teer, & Rubenstein, 1978; Morrison, Ellis, & Chappel, 1997; Gilhooly & Gilhooly, 1980). AoA has been shown to correlate with other variables, more notably word familiarity and frequency; however, factorial studies have indicated that it has a unique weight in various processes (see for example Burani, Arduino, & Barca, 2007).

This variable, obtained from empirical ratings, may be a strong predictor of adult performance in tasks such as auditory and visual lexical decision, word naming, picture naming, object naming and face recognition (Barca, Burani, & Arduino, 2002; Bates, Burani, D'Amico, & Barca, 2001; Brown & Watson, 1987; Burani et al., 2007; Carroll & White, 1973; Colombo & Burani, 2002; Ellis & Lambon Ralph, 2000; Ellis & Morrison, 1998; Gerhand & Barry, 1999; Moore & Valentine, 1998; Morrison & Ellis, 1995; Turner, Valentine, & Ellis, 1998). Typically, early acquired words are recognized and read faster than late acquired words, and faces that are early «acquired» and recognized in childhood are recognized faster than late acquired faces.

Although potentially relevant, AoA has never been considered in idiom studies. Therefore, we decided to include this variable in the corpus. The data will be particularly useful in work with children, where it is well-

known that mastering idiomatic language is an ability that is acquired slowly and late in development (Ackerman, 1982; Levorato & Cacciari, 1992, 1995; Levorato, Nesi, & Cacciari, 2004; Nippold & Martin, 1989).

Predictability

With respect to idioms, predictability is defined as the probability of completing an incomplete string idiomatically. For example, “Mario è al settimo ...”/“Mario is in seventh...” is typically completed by speakers idiomatically, with *cielo/heaven*, even though other, literal completions are possible (e.g., *posto/place*). In contrast, “Mario ha le mani ...”/“Mario has his hands ...” is completed by most speakers literally (e.g., *sporche/dirty*), rather than idiomatically (*bucate/pierced*; “be a spendthrift”).

In general, predictability is a property of sentences or phrases that may influence lexical processing. It has been shown that the processing of an upcoming word can be facilitated in a predictable context relative to a less predictable one (Erlich & Rayner, 1981; Rayner & Pollatsek, 1989; Schwanenfluegel & Shoben, 1985; Zola, 1984).

Cacciari and Tabossi (1988) showed that the figurative meaning of highly predictable idioms is recovered sooner than that of low predictable ones. Similar results were obtained by Titone and Connine (1994a). The effects of predictability on idiom processing have also been reported in several other studies (Tabossi & Zardon, 1993, 1995; Tabossi, Fanari, & Wolf, 2008).

Predictability is particularly relevant in the framework of the Configuration Hypothesis of idiom recognition (Cacciari & Tabossi, 1988). According to this view, idioms are not represented as individual items, but as configurations of words. These words are the same that are represented in the lexicon and processed during literal language comprehension. However, speakers know that some sequences of words, e.g., *kick* followed by *the bucket*, have a figurative meaning (e.g., “to die”), and upon identifying a sequence of this sort, the figurative meaning associated with it is retrieved from memory. During speech comprehension, an idiomatic string is processed word by word, just like any other piece of language, until enough information has accumulated to render the sequence of words identifiable as – or highly expected to be – a memorized idiom. Without a context biasing the figurative interpretation, the string's idiomatic meaning is activated only when it becomes predictable and is recognized as an idiom.

Syntactic flexibility

Idioms cannot enter all the syntactic operations without losing their figurative interpretation. However, there is a

large degree of variability among expressions, and while some are almost entirely frozen (e.g., *kick the bucket*), others accept almost every operation (e.g., *lay down the law*). Understanding the syntactic behavior of idioms is a central task for any theory of idiom processing, and several alternatives are currently under debate (Gibbs & Nayak, 1989; Sprenger et al., 2006; Tabossi, Fanari, & Wolf, 2009; Tabossi, Wolf, & Koterle, 2009). So far, there are no data bases available on idiom flexibility. However, in the study of idiom syntax, knowing the syntactic behavior of individual idiomatic expressions is a valuable tool that can be used to manipulate the syntactic variable experimentally (Tabossi et al., 2008). For this reason, we chose to collect ratings on the flexibility of the idioms in our corpus.

In English, the operations usually considered are adverb insertion, adjective insertion, present participle, action nominalization, and passive (Gibbs & Nayak, 1989). In Italian, present participle and action nominalization are not possible. In their place, we used two other operations: left dislocation and movement. In left dislocation a nominal constituent (e.g. *ghiaccio/ice* in *Mario rompe il ghiaccio/ Mario broke the ice*) is moved to the left, and it is typically substituted in its original position by a clitic pronoun (e.g. *Il ghiaccio, Mario lo rompe in breve tempo/ The ice, Mario broke it in a short time*; Renzi, Salvi, & Cardinaletti, 1988). Movement operations have the effect of moving constituents (usually one at the time) from one part of a phrase-maker to another as in interrogative and exclamative sentences. For instance *Franco trovò la chiave/Franco found the key* becomes *Che chiave ha trovato Franco?/What key did Franco find?*

Both operations have been used to explore the syntax of Italian idioms, and are known to have a rather limited use (Bianchi, 1993). Typically, idioms that do not accept most operations are considered nonflexible, whereas idioms that accept most operations are considered flexible. Fraser (1970) proposed a hierarchy in idioms' acceptance of syntactic operations. However, his proposal has never received empirical support, and at present we do not know whether there is any correlation in idioms' ability to accept syntactic alterations. In other words, we do not know whether an idiom that, say, does not accept movement is also likely not to accept the passive, or whether acceptability for the two operations is independent from each other. In order to allow researchers to consider these alternatives, in our corpus, in addition to presenting for each idiom its mean rating of general flexibility obtained by averaging across the ratings for the different syntactic operations for an idiom, we decided to include its rating of acceptability for each of the operations.

Literality

Literality refers to the degree to which an idiomatic expression has a plausible literal interpretation (Popiel & McRae, 1988). For example, *break the ice* is a literal idiom as it has both a figurative and a literal interpretation. However, neither *go bananas* nor *shoot the breeze* have a literal interpretation, although for different reasons: the former is syntactically ill-formed as *go* is an intransitive verb that cannot normally take a direct complement; the latter is syntactically acceptable, but semantically anomalous as the breeze is not the kind of object that can normally be shot.

According to the traditional view, English idioms with no literal counterpart are less syntactically flexible than idioms that have a literal counterpart (Fraser, 1970). In contrast, Gibbs and Nayak (1989) argued that non literal idioms are perceived as more decomposable and hence more flexible than literal idioms (see below for compositionality).

So far, the evidence on the role of literality on idiom processing has shown inconsistent results. Brannon (Popiel & McRae, 1988) found longer classification times for literal idioms compared to idioms that had only a figurative interpretation. Opposite results were obtained by other authors in a variety of tasks (Cronk & Schweigert, 1992; Estill & Kemper, 1982; Mueller & Gibbs, 1987; Schweigert, 1986; Swinney & Cutler, 1979). In an incidental-memory study conducted by Botelho da Silva and Cutler (1993), participants recalled idioms much better than control strings, but there was no reliable difference in their ability to remember literal and non literal idioms. In a cross-modal study, Titone and Connine (1994a) explored the activation of the literal meaning of the final word of idiomatic expressions as a function of the predictability and literality of the strings. They found that with predictable expressions the meaning of an idiom's last word is activated with literal strings, but is not activated with non literal ones. The role of literality is still in need of clarification.

Compositionality

One central principle in “literal” language is compositionality, which dates back to Frege (1892). Consider “The young boy ate the apple”. The meaning of the noun phrase *The young boy* is obtained by combining the meanings of *the*, *young*, and *boy* according to the grammatical analysis of the noun phrase, the meaning of the verb phrase *ate the apple* is obtained by combining *ate* and *the apple* according to the grammar of the language. Finally, the meaning of the whole sentence is obtained by combining the meaning of the noun phrase with the

meaning of the verb phrase, again in accordance with the grammar of the language.

For a long time, idiomatic expressions have been considered a typical example of non compositionality (Chomsky, 1980; Fraser, 1970; Katz, 1973). However, in contrast with the view prevailing at the time, Nunberg (1978) proposed that several intriguing grammatical phenomena shown by idiomatic expressions may be accounted for on the basis of their semantic properties. He introduced the notion of idioms as combining expressions. Compositionality here refers to the fact that the constituents of some idioms “carry identifiable parts of the idiomatic meaning” (Nunberg et al., 1994, p. 496). In *pop the question*, for example, there is a clear correspondence between *pop* and *question* and the relevant parts of the figurative meaning “propose marriage”. In contrast, in *kick the bucket* there is no sense in which the figurative meaning “die” is distributed over *kick*, *the*, and *bucket*. Accordingly, *pop the question* is a decomposable idiom, and *kick the bucket* is not. Nunberg et al. (1994) argued that the more decomposable an idiom is, the more likely it is to be syntactically flexible.

Compositionality was introduced in psycholinguistics by Gibbs and colleagues, who, based on this notion, developed the Idiom Decomposition Hypothesis (Gibbs, Nayak, & Cutting, 1989). According to this view, idioms are at least partly decomposable and speakers have shared intuitions about how the meanings of the parts contribute to the idiomatic meaning. These intuitions determine the syntactic behavior of the idioms: the more an idiom is seen as analyzable, the more likely that idiom will be treated as syntactically flexible (Gibbs & Nayak, 1989). Moreover, idiom processing is based on compositionality: semantically decomposable expressions may be analyzed compositionally. Each component is retrieved from the mental lexicon and combined with the other components of the string according to their syntactic relations. In contrast, the meaning of nondecomposable idioms is retrieved directly from the lexicon. Thus, the comprehension of decomposable and nondecomposable idioms is achieved through different processes: decomposable idioms involve the same mechanisms of lexical retrieval and syntactic parsing taking place during the comprehension of literal expressions, whereas nondecomposable idioms require processes akin to those operating in the recognition of individual words (Gibbs, Nayak, Bolton et al., 1989).

The empirical adequacy of the Idiom Decomposition Hypothesis is debatable, and several studies have shown data inconsistent with its predictions (Cutting & Bock, 1997; Tabossi et al., 2008; Titone & Coninne 1994a, 1999). However, the notion has become a central one in the field, and controlling this dimension is crucial in any study on idiom comprehension or production.

Methods

Participants

A total of 740 northern Italian speakers (380 female and 360 male), ranging from 17 to 50 years of age (mean age: 26;4 - SD: 3;5), volunteered in this study. They had 13 years of school or more and were high school students and teachers and undergraduate university students. These participants allowed us to have a minimum of 40 ratings for each idiom and each of the variables considered.

Material

The idiomatic expressions have been selected from the *Dizionario Rizzoli dei modi di dire della lingua italiana* (Rizzoli Dictionary of the ways of saying things in the Italian language) (Quartu, 1993). Three Italian speakers from different linguistic areas selected idioms, using the following criteria: all the expressions were verbal expressions, i.e. expressions such as *by and large* were not selected; according to the three speakers' intuition, the idioms were known in all the country, and not restricted to a regional use. Half of the expressions were short (i.e., they contained only a content word after the verb; e.g., *alzare il gomito/raise the elbow*, “drink too much alcohol”), and half were long (i.e., they contained more than a content word after the verb; e.g., *essere al settimo cielo/be on seventh heaven*). Also, judges selected idioms so that, according to their intuition, half were literal and half were not. Although pre-selected, literality has been investigated in the study.

Database

For each of the 245 selected idioms, the following values are reported: length, knowledge, familiarity, AoA, predictability, syntactic flexibility, literality and compositionality. In accordance with common use in the literature, length is given in terms of number of words in the string (Fanari et al., 2010). All measures except length were obtained from empirical ratings given by Italian speakers.

For familiarity, AoA, syntactic flexibility, literality and compositionality the mean values and standard deviations (SDs) are reported for each expression. For predictability the values reported is the proportion of participants who completed the initial fragment of an expression (e.g., *be in seventh...*) idiomatically (e.g., *heaven*). For meaning knowledge the database reports, for each idiom, the proportion of participants knowing the expression (see below for detail about the data collecting procedure).

Procedure

Knowledge scale

The corpus has been divided into two lists of 122 and 123 idioms, respectively, balanced for length and literality (see above). A total of 88 participants took part in this ratings (44 for each list). Each participant received a booklet with the idiomatic expressions, preceded by the following instructions:

For each of the idiomatic expressions you will see, explain its meaning, if you know it. Do nothing otherwise. To express the meaning, you may use a single word, a sentence, or two or more sentences. For example, for *avere il pollice verde/have the green thumb* you may say something like “obtain very good results in gardening.”

Familiarity scale

The aim of the familiarity test was to establish the degree of familiarity of each idiomatic expression relative to the other idiomatic expressions. Therefore, in agreement with the literature, no literal fillers were used (e.g. Coninne, Blasko, Brandt, & Layer, 1992; Cronk & Schweigert, 1992; Libben & Titone, 2008; Titone & Coninne, 1994b).

A total of 42 participants received a booklet containing all the 245 expressions, each paired with a 7-point scale. For each participant, idioms were presented in a different random order. Participants were given the following instructions:

For each idiom in the following pages you need to judge how well you think the expression is known by people like you, independently from whether or not you know it. To express your judgment, you have to use the scale next to each expression: 1 is for expressions you think are not known at all, and 7 is for expressions you think are known to everybody. Please use intermediate values for intermediate degrees of knowledge. Please mark clearly each of the point you choose.

A few examples were also provided.

AoA scale

All the idioms were presented on a booklet to 40 participants, who were requested to evaluate on a 7-point scale at what age, in their opinion, they had learned

each idiomatic expressions. They received the following instructions:

“In this study we intend to collect your evaluations on the age at which you think you have learnt the meaning of a series of idiomatic expressions, whether in spoken or in written form. For each of the idioms on the following pages you will find a scale like the one below: 0–4 years (1), 5–6 years (2), 7–8 years (3), 9–10 years (4), 11–12 years (5), 13–14 years (6), 15 years or later (7).

Your task is to mark, for each idiom, the value corresponding at when you think you learned the expression. For example, if you think you learned an expression at the age of 7 years, mark the value 3 in the scale; if you think you have learned the expression at 4 years, mark the value 1. Try and use all the values in the scale. Give only one mark to each expression”.

Predictability scale

Each idiom has been inserted in a short sentence. Sentence frames prior to the idioms were as neutral as possible with respect to a literal or figurative continuation (e.g., for *trovare la chiave/find the key* “find the solution”, the frame was: “Antonella ha trovato la _____”/“Antonella found the _____”). For each idiomatic sentence, two literal filler sentences were created changing the idiomatic verb with another verb matched with it for frequency (e.g., “Antonella ha rifiutato la _____”/“Antonella rejected the _____”; “Antonella ha ricamato la _____”/“Antonella embroidered the _____”).

From a total of 735 sentential frames five lists were created. Each list contained 49 idiomatic sentences and 98 literal filler sentences semi-randomly selected, with the following constraints: about half of the sentences were long and half were short, and within a list, idiomatic sentences and corresponding literal sentences created from them never occurred. Sentences in each list occurred in two different random orders.

A total of 210 Italian speakers participated in this rating. Five groups of 42 participants each were randomly assigned to one list. Within each group, 21 participants received one list in one random order, and 21 received the same list in the other random order. Each participant received a booklet containing one list of incomplete sentences, preceded by the following instructions:

In the following pages you will find incomplete sentences. Your task is to complete the sentence with one word only. Make sure that the resulting sentence makes sense. Write your answer next to the sentence

fragment. For example, if you see a fragment like “Maria desidera molto un ____/”Maria longs for a ____”, you may decide to complete the sentence with *cane/dog*, in which case you write down *cane* next to the sentence. However, if the first word that occurs to you is *piede/foot*, you have to pause and think to another one, because “Maria desidera molto un piede/”Maria longs for a foot” does not make sense.

Syntactic flexibility scale

Following Gibbs and Gonzales (1985), each idiomatic expression was inserted in a sentence. In each sentence the idiomatic expression occurred with one of the following syntactic modifications:

1. adverb insertion
2. adjective insertion
3. left dislocation
4. passive
5. movement

Due to syntactic limitations, not all idioms entered all the operations. In particular, idioms with intransitive or reflexive verbs were not passivised. Overall, 1132 sentences were created. Each idiomatic sentence was paired with its paraphrase. For example, in the case of adverb insertion, the idiom *spezzare una lancia/break a lance*, “defend” occurred in the sentence “Mario spezza spesso una lancia in favore di Giulia quando gli chiedono un’opinione/”Mario often breaks a lance in favour of Giulia when they ask him for an opinion”, and was paired with the paraphrase “Mario prende spesso le difese di Giulia quando gli chiedono un’opinione/”Mario often defends Giulia when they ask him for an opinion”. Table 1 presents an example of two idioms in each of the five syntactic forms.

From the sentence pairs thus obtained five lists were created. Four lists contained 226 pairs and one list contained 228 pairs. A single idiom appeared only in one list, and each list contained approximately the same number of operations. Each list appeared in two different random orders.

A total of 200 Italian speakers took part in this rating. An equal number of them (40) was randomly assigned to each list. Within each list, 20 participants received one random order, and 20 received the other order. They were requested to evaluate on a 7-point scale to what extent the meaning of an idiomatic sentence was similar to the meaning of the paraphrase. It is assumed that ratings reflect the degree to which an idiom can be syntactically altered and still maintain its figurative

meaning. Participants were given a booklet containing one list, preceded the following instructions:

In the following pages you will find a series of idiomatic phrases paired with their paraphrases. Next to them, you will find a 7-point scale. Your task is to judge, for each of these pairs, to what extent the meaning of idiomatic sentence is similar to the meaning of the paraphrase. It is important that you give your judgment keeping in mind the form in which the idiomatic sentence is presented, without thinking to what it might be, if it were presented in a different form. Assign the scores on the scale: 1 means not at all similar and 7 means very similar. Use intermediate scores for intermediate degrees of similarity.

Literality scale

Forty participants judged for each of the 245 idiomatic expressions whether or not it had a literal interpretation, using a 7-point scale, ranging from *it has no literal meaning* (1) to *it has a clear literal meaning* (7). Participants received a booklet containing all the expressions in a random order, each paired with the seven-point scale. They were given the following instructions:

For each of the idioms in the following pages, you will have to give a judgment concerning its literality. All idioms have a figurative meaning, but only some of them have, in addition, a literal meaning. For example, *tenere alta la bandiera/hold the flag high* figuratively means “honour one’s own cause”, whereas literally it means “keep the flag lifted up”. In contrast, *restare di ghiaccio/remain of ice* figuratively means “be astounded”, but it is not clear what it could mean literally. Your task is to decide for each expression whether it has a literal interpretation on a 7-point scale: 1 indicates that the expression does not have a literal interpretation, and 7 indicates that it has a clear literal interpretation. The intermediate points indicate intermediate intuitions concerning the plausibility of a literal interpretation. Please make a clear mark on the selected point.

Compositionality scale

Following Gibbs, Nayak, and Cutting (1989) and Gibbs and Nayak (1989), each idiomatic expression was paired with a paraphrase of its figurative meaning (e.g., Idiom: *Prendere il toro per le corna/take the bull by the horns*; Paraphrase: “Affrontare una situazione direttamente/”Face a situation

Table 1 Example of different syntactic operations

Alzare il gomito/raise the elbow, “drink too much (alcohol)”

Adverb Insertion

Piero alza sempre il gomito quando va a cena da Teresa/Piero raises always the elbow when he has dinner at Teresa’s

Adjective Insertion

Piero alzò il solito gomito quando andò a cena da Teresa/Piero raised the usual elbow when he had dinner at Teresa’s.

Left Dislocation

Il gomito Piero lo alza quando esce con Giovanni/The elbow Piero raises it when he goes out with Giovanni.

Passive

Il gomito fu alzato da Piero la sera che uscì con Giovanni/The elbow was raised by Piero the night he went out with Giovanni.

Movement

Che gomito ha alzato Piero quando è andato alla festa di Teresa!/What an elbow Piero raised when he went to Teresa’s party!

Battere il ferro finché è caldo/hammer the iron while it is hot, “take advantage of the occasions while are propitious; do something when things are going well for you and you have a good chance to succeed”

Adverb Insertion

Franco ha battuto abilmente il ferro finché era caldo per ottenere la promozione/Franco cleverly hammered the iron while it was hot in order to get the advancement.

Adjective Insertion

Franco ha battuto tutto il ferro finché era caldo per ottenere la promozione/Franco hammered all the iron while it was hot in order to get the advancement.

Left Dislocation

Il ferro Franco lo batte finché è caldo ogni volta che si presenta l’occasione/The iron Franco hammers it while it is hot, every time that he has the possibility to do so.

Passive

Il ferro è stato battuto da Franco finché era caldo per ottenere quella promozione/To get his advancement the iron has been hammered by Franco while it was hot.

Movement

Che ferro ha battuto Franco finché era caldo per ottenere quella promozione?/Which iron did Franco hammer while it was hot in order to get the advancement?

in a straight way”). The pairs so obtained were divided into three lists: two lists of 82 pairs and one list of 81. Within each list pairs occurred in a random order.

A total of 120 Italian speakers took part in this rating, 40 for each list. They were given a booklet containing one list, preceded by the following instructions:

In the following pages you will find a list of idiomatic expressions, each paired with the paraphrase of its figurative meaning. For example, *essere in vena/be in vein* may occur paired with “sentirsi propenso a fare qualcosa”/“to be in the mood for something”. Next to each pair you will also find a 7-point scale: 1 means *not at all* and 7 means *completely*. For each idiom/paraphrase pair, you must select and mark one value in the scale, depending on how decomposable the idiom is. An idiom is decomposable if its constituent parts contribute to the meaning of the expression. For example, the paraphrase of *abbassare la cresta/lower the crest* is “ridurre le proprie pretese”/“reduce one’s pretensions”. This idiom is considered decomposable because there is an intuitive relation between *abbassare/lower* and *ridurre/reduce*, and between *la cresta/*

the crest and *le proprie pretese/one’s pretensions*. In contrast, the paraphrase of *essere in vena/be in vein* is “sentirsi propenso a fare qualcosa”/“to be in the mood for something”, and there is no intuitive relation between the parts of the idiom and the parts of the paraphrase. This idiom is considered nondecomposable. Remember to use the intermediate values for idioms that in your intuition are partly decomposable.

Results and discussion

Length is expressed in number of words. For literality, familiarity, AoA, compositionality we calculated the mean values and SDs obtained by each item for all the participants. For syntactic flexibility we calculated the mean values and SDs obtained by each idiom for each of the operations it underwent. Moreover, we calculated the mean degree of flexibility (and SDs), averaging the mean values obtained by each idiom on the single operations. For knowledge and predictability, we calculated the proportion of correct definitions or idiomatic completions obtained by

each idiom, giving 1 to correct responses and 0 to incorrect or missing responses. These values for all the idioms are reported in the Excel file in the supplemental material. In Table 2 we reported mean, standard deviation and range of data for each dimension.

Length

Idioms were pre-selected so that about half of them had only one content word after the verb. The corpus length range, in words, is 2–7. For the entire corpus the mean length in words is 3.9 (SD 1.1). For the 119 idioms with only one content word after the verb the mean is 2.9 (SD 0.3), for the 126 longer idioms is 4.9 (SD 0.8).

Knowledge

Twenty-five expressions are perfectly known (i.e., score of 1.0). An example is *voltare pagina/turn page* (“change lifestyle”). The least known expression is *cascare dal pero/fall down from a pear tree* (“become painfully aware of a sad state of affairs”) with a score of 0.1. The mean proportion of correct definitions in the corpus is 0.8 (SD 0.2). An example of idiomatic expression moderately known (i.e., known by half of the participants) is *fare un colpo/make a stroke* (“commit a robbery”), with a score of 0.50.

On average, knowledge is rather high. This is very unlikely to indicate that in general idioms are well known to speakers. More probably, the high score reflects an initial bias, where the judges mostly selected idioms known to them. Clearly, many rare idioms are not included in the corpus. Nonetheless, the range is very wide, showing large differences in people’s knowledge of idiomatic expressions.

Familiarity

Avere il mestolo in mano/ have the dipper in hand (“detain the power”) received the lowest value of familiarity (3.4,

SD 2.4), and *essere al settimo cielo/be in seventh heaven* received the highest (6.8, SD 0.5). An example of an idiom of intermediate familiarity is *spaccare il capello in quattro/ chop the hair in four* (“examine something with too much pedantry”): score 4.0, SD 2.2. The mean familiarity rating is 5.7 (SD 0.7).

As for knowledge, the high mean score we obtained probably reflects a bias in the initial selection of the expressions. It should be noted, however, that work on idiom familiarity uses expressions from high to low familiarity among known expressions, and the corpus does offer a range of this sort.

AoA

In our database *firmare una cambiale in bianco/sign a bill of exchange in white* (“trust somebody too much”) is rated as the last idiom to be acquired (6.0, SD 1.0, 13/14 years), *contare le pecore/ count the sheeps* (“try to fall asleep”) as the first (2.2, SD 1.5, 5/6 years). The mean AoA for the entire corpus is 4.1 (SD 0.7, 9/10 years).

These entirely new data are consistent with the experimental literature, suggesting that these expressions are acquired late, sometimes not until adolescence (Ackerman, 1982; Levorato & Cacciari, 1992, 1995; Levorato et al., 2004; Nippold & Martin, 1989).

Predictability

Consistent with previous studies (Libben & Titone, 2008; Titone & Connine, 1994b), we found a wide range of variability in predictability (from 0.0 to 1.0). In other words, some idioms are never predicted (e.g., *lasciare il campo/leave the field*, “give up the undertaking”), and others are always predicted (e.g., *pendere dalle labbra di qualcuno/hang on somebody’s lips*, “hang on someone’s every word”). The mean proportion of idiomatic completion is 0.5 (SD 0.3). An example of a string that is completed idiomatically 50% of times is *trovare tutte le porte aperte/find open all the doors*, “find no obstacles” (0.5, SD 0.5).

Table 2 Range, mean and standard deviation for all the idiom dimensions: idiom Length (LEN), Knowledge (KNOW), Familiarity (FAM), Age of Acquisition (AoA), Predictability (PRED), Syntactic Flexibility (FLEX), Literality (LIT), Compositionality (COM)

	LEN	KNOW	FAM	AoA	PRED	FLEX	LIT	COM
Range	2–7	0.1–1.0	3.4–6.8	2.2 (5/6 years)–6.0 (13/14 years)	0.0–1.0	3.4–6.1	1.1–6.9	1.55–6.6
Mean	3.9	0.8	5.7	4.1 (9/10 years)	0.5	5.1	4.1	4.1
SD	1.1	0.2	0.7	0.7	0.3	0.5	2.0	1.0

Syntactic flexibility

The overall mean flexibility ratings range from 3.4 (SD 2.2) for *andare in giro/go in round*, “wander” to 6.10 (SD 1.6) for *essere al settimo cielo/be in seventh heaven*. An example of an idiom that was judged with an intermediate degree of flexibility is *essere in ballo/be in ball*, “be involved in something without the possibility to avoid it”: 4.0 (SD 2.1). The mean rating for the entire corpus is 5.1 (SD 0.5).

The mean acceptability values and SDs for the single syntactic operations considered in this study are reported in Table 3.

For adverb insertion, the mean rating ranges from 3.5 (SD 2.1) for *cadere una tegola in testa/fall a tile on the head*, “suddenly happening of something bad” to 6.6 (SD 1.4) for *dare alla luce/give to the light*, “give birth”. An intermediate value was obtained for *cascare dal pero/fall from the pear tree*, “become painfully aware of a sad state of affairs” (M 4.0, SD 2.16).

For adjective insertion, values range from 3.1 (SD 2.0) for *prendere per il collo/take by the neck*, “compel” to 6.5 (SD 1.1) for *mettere la mano sul fuoco/put the hand on the fire*, “be absolutely sure”. An intermediate value was obtained for *spremersi le meningi/squeeze his/her own meninges*, “rack one’s brains” (M 4.0; SD 2.2).

The mean flexibility value for left dislocation ranges from 3.3 (SD 2.1) for *covare una serpe in seno/brood a viper in breast*, “do good to someone who represents a potential danger” to 6.3 (SD 1.4) for *essere due gocce d’acqua/be two water drops* “be exactly alike”. *Cascare dal pero/fall from the pear tree*, “become painfully aware of a sad state of affairs” obtained an intermediate value M 4.0; SD 2.1).

Values for the passive operation range from 2.8 (SD 1.9) for *coltivare il proprio orticello/cultivate oneself own little vegetable garden*, “make only your own business” to 6.7 (SD 0.7) for *fare un lavoro coi fiocchi/do a job with the bow ties*, “do an excellent job”. An intermediate value was obtained by *rischiare le penne/risk the feathers*, “risk the life” (M 4.0; SD 2.5).

The last syntactic operation, movement, showed mean scores from 2.6 (SD 2.1) for *prendere per il collo/take by the neck*, “compel” to 6.5 (SD 0.9) for *gettare il denaro*

dalla finestra/ throw the money out the window, “waste money”. *Tagliare la corda/cut the rope*, “get away, escape” had an intermediate rating: M 4.0 (SD 2.1).

Literality

The scores for literality range from 1.1 (SD 0.2) (i.e., *Pendere dalle labbra di qualcuno/ hang on somebody’s lips*, “hang on someone’s every word” to 6.90 (SD 0.30) for *voltare paginai/turn page*, “change lifestyle”. A moderately literal idiom is *buttarsi dietro le spalle/ throw behind the shoulders* (“forget”), whose score is 4.0, SD: 2.40. The mean literality rating is 4.1 (SD 2.0).

These results are very similar to those reported by Libben and Titone (2008), who used a 5-point scale, but whose mean literality value, once transformed on a 7-point scale, is 4.2. Somewhat higher is the mean value in Titone and Coninne (1994b), whose value corresponds to 5.6 in a 7-point scale.

Compositionality

The lowest score of compositionality was 1.6 (SD 1.28) for *mangiare la foglia/eat the leaf*, “realize what is going on”, whereas the highest was 6.6 (SD 0.68) for *essere ad un bivio/to be at the crossroads*, “be in the situation in which a critical decision must be made”. An example of an idiom with an intermediate score of compositionality is *andare a genio/go to genius*, “be pleasant; be the one’s liking”: 4.0, SD 2.25. The compositionality ratings were normally distributed (mean 4.1, SD 1.0; Shapiro-Wilk $W = .99$, $p = .31$; Skeweness = $-.01$).

To explore for how many idioms speakers’ intuitions on compositionality agreed we used the 67% agreement criterion employed in prior studies (Titone & Connine, 1994b): by this criterion only 18% of idioms were consistently rated as decomposable, and 10% of the expressions were consistently rated as nondecomposable (i.e., had a score of 4 or less on a 7-point scale). Clearly, for most of the 245 idiomatic expressions judgments were not consistent. Although in contrast with Gibbs and Nayak (1989), for decomposable idioms our corpus is consistent with Titone and Connine (1994b), who found that 22% of idioms were rated as decomposable. However they also found a tendency to evaluate strings as nondecomposable (42%). In fact we obtained only 10% agreement for the nondecomposable expressions. This inconsistency is disturbing, as all the studies used the same procedure. Perhaps, the differences reflect the instability of the results with respect to this value (Tabossi et al., 2008).

Table 3 Mean acceptability values and sds obtained for the individual syntactic operations

Syntactic Operation	Mean	SD
Adverb insertion ($n = 245$)	5.5	0.6
Adjective insertion ($n = 244$)	5.0	0.7
Left dislocation ($n = 245$)	5.1	0.6
Passive ($n = 157$)	5.1	0.7
Movement ($n = 241$)	4.6	0.7

Relations among the variables

To assess whether there are relations among the variables considered, we conducted a correlation on the following variables: length, knowledge, familiarity, AoA, predictability, syntactic flexibility, literality and compositionality. For syntactic flexibility we used the mean scores obtained by averaging across the ratings for the different syntactic operation for an idiom. The correlation matrix is presented in Table 4

To reduce the probability to incur a Type I error, and to consider as reliable only correlations that capture at least 4% of variance, we set the level of significance equal to .001.

The strongest correlation is between length and predictability ($r = 0.59$; $N = 245$; $p < .001$), capturing 35% of variance: the longer an idiomatic fragment, the more likely it is to be completed idiomatically. This is consistent with Fanari et al. (2010), who found that without a biasing context the figurative meaning of an idiomatic expression is available at the offset of long idioms (whether predictable or not), but it is not still available at the offset of short idioms.

Familiarity shows a strong inverse correlation with AoA ($r = -0.58$, $N = 245$, $p < 0.001$), accounting for 29.00 % of variance. The data suggest that idioms that are familiar are likely to be the first to be made available in the linguistic environment and, thus they are also the first acquired by children.

Unsurprisingly, familiarity also correlates strongly with knowledge ($r = 0.49$, $N = 245$; $p < .001$, 24.01% variance), suggesting that in general speakers know better the meaning of an idiom they are familiar with. This is consistent with the correlation found by Titone and Connine (1994b), even though the instructions for collecting knowledge differed somewhat in their study and in the present one.

The two other correlations capturing more than 10% variance are knowledge and compositionality ($r = 0.35$, $N = 245$, $p < 0.001$, 12.00% variance) and familiarity and literality ($r = -0.32$, $N = 245$, $p < 0.001$, 10.24% variance). The former indicates that the better one knows the meaning

of an idiom, the more s/he is likely to see it as compositional. This may suggest that a better knowledge might help people to develop intuitions about the individual contributions that a string's parts make to the overall figurative meaning of the expression. Familiarity inversely correlates with literality: The more one is familiar with the meaning of an idiom, the less likely s/he is to treat it as a literal expression. Perhaps, this indicates that for familiar idioms the figurative meaning is stronger than the possible literal counterpart.

All the other correlations explain less than 10% of variability, and hence do not indicate very strong relations among the variables. However, it is interesting to observe that, albeit weak, some of these correlations are consistently observed. In particular the correlation between familiarity and predictability ($r = 0.21$, $N = 245$, $p = 0.001$, 4.00%) replicates Titone and Connine (1994), who reported a reliable positive correlation between predictability and both the measures they used to assess familiarity, namely frequency and meaningfulness. A similar effect was also found by Libben and Titone (2008). Likewise, the correlation between compositionality and flexibility ($r = 0.28$, $N = 245$, $p < .001$, 7.00% variance) replicates a similar result reported by Gibbs and Nayak (1989). In addition, in the present corpus knowledge correlates negatively with AoA ($r = -0.29$, $N = 245$, $p < .001$, 8.00% variance) and familiarity correlates with flexibility ($r = 0.28$, $N = 245$, $p = .001$, 7.00% variance).

To our knowledge, there is no evidence on whether the flexibility of an idiom in one operation is correlated with its flexibility in other operations. Fraser (1970) proposed that phrasal idioms should be considered as a variety of monomorphemic lexical entries, which permit certain types of syntactic operations. These operations are in a hierarchy such that if an idiom belongs to a level of the hierarchy, it will accept the operation that characterises that level. Moreover, it will also accept automatically all the operations specified at lower levels: the higher up on the hierarchy, the more syntactically unfrozen the idiom. For example, at one level are idioms that accept passive (e.g., *blow the whistle on*; *pop the question*; *spill the beans*).

Table 4 Correlation matrix for Length (LEN), Knowledge (KNOW), Familiarity (FAM), Age of Acquisition (AoA), Predictability (PRED), Syntactic Flexibility (FLEX), Literality (LIT), Compositionality (COM)

	LEN	KNOW	FAM	AoA	PRED	FLEX	LIT	COM
LEN	1.00	-0.03	0.03	0.14	0.59*	-0.01	0.04	0.05
KNOW		1.00	0.49*	-0.29*	0.02	0.45*	-0.17	0.35*
FAM			1.00	-0.58*	0.21*	0.28*	-0.32*	0.09
AoA				1.00	-0.06	-0.11	0.19	-0.11
PRED					1.00	0.01	0.04	0.05
FLEX						1.00	-0.17	0.28*
LIT							1.00	0.18
COM								1.00

* $p < 0.001$

Table 5 Correlation matrix of the acceptability scores for Adverb Insertion (ADV), Adjective Insertion (ADJ), Left Dislocation (LDIS), Passive (PASS) and Movement (MOV). The Correlations Have Been Calculated over the 154 Idioms that Accepted All the Operations

	ADV	ADJ	LDIS	PASS	MOV
ADV	1.00	0.47*	0.58*	0.45*	0.23
ADJ		1.00	0.29*	0.35*	0.41*
LDIS			1.00	0.50*	0.35*
PASS				1.00	0.31*
MOV					1.00

* $p < 0.001$

Lower in the hierarchy there are idioms that accept adverb insertion (e.g., *give ground to*; *drop a line to*; *do a good turn to*). Because the former group of idioms is higher in the hierarchy, in addition to accepting passive, it will also accept adverb insertion, whereas the reverse is not true for the latter group of expressions. However, as mentioned before, Fraser's notion of an idiom hierarchy has never found empirical support in psychological research (Botelho da Silva & Cutler, 1993; Gibbs & Gonzales, 1985; Tabossi et al., 2008), and the only point on which the literature agrees is that adverb insertion is the most easily accepted among the various operations (Gibbs & Gonzales, 1985; Connine, Blasko, Brandt, & Kaplan, 1992). Accordingly, we decided to explore whether there are correlations among the acceptability scores that idioms obtain in the different operations. We used the same criteria as for the former correlations.

All the operations, except adverb insertion and movement, were reliably correlated to one another (Table 5). However, adjective insertion and dislocation, and passive and movement accounted for less than 10% of variance. The strongest correlations, explaining 33.64 % and 25.00 % of variance respectively, were between left dislocation and adverb insertion ($r = 0.58$, $N = 154$, $p < .001$) and left dislocation and passive ($r = 0.50$, $N = 154$, $p < .001$). The variance captured by the remaining correlation ranged between 22.00% (adverb and adjective insertion) and 12.25% (adjective insertion and passive, and left dislocation and passive).

This is novel evidence, and may be of help to researchers who wish to generalize from results on a single operation to syntactic flexibility in general.

Finally, we explored whether the different syntactic manipulations are related with the other variables considered in this study except length. (There is no theoretical reason to posit that the length of an idiomatic expression might have any relation with its syntactic behavior.) To this end, we computed additional correlations (Table 6). The results showed that only two correlations explained at least 10% variance: Knowledge correlated positively with both adverb insertion ($r = 0.42$, $N = 245$, $p < .001$, 18% variance) and left dislocation ($r = 0.41$, $N = 245$, $p < .001$, 17% variance). Although statistically robust, it is difficult to imagine a theoretical reason why knowledge, but not familiarity, is related specifically to these two operations. This is particularly surprising, since adverb insertion and left dislocation are rather different in terms of difficulty, the former being the easiest of all the manipulations. Perhaps equally surprising is the lack of evidence supporting a relation between the syntactic operations and AoA: in principle, one could imagine that people are more likely to accept difficult operations in early acquired, and thus best known, expressions. Also noteworthy is the fact that the correlations between compositionality and the single operations capture a very low degree of variance, ranging from 8% to 1.4%. This corroborates the view that the relation between flexibility and compositionality, if it exists, is very weak indeed.

Conclusions

The main aim of this study was to present a database of Italian verbal idiomatic expressions with variables that are known to influence the processing of these expressions. Norms for length, knowledge, familiarity, AoA, predictability, syntactic flexibility, literality and compositionality were provided. Similar instruments are available for English, but this is a new instrument for Italian. Also, ratings for AoA and syntactic flexibility are made available for the first time.

Table 6 Correlation values for the individual syntactic operations and the other idiom dimensions: Knowledge (KNOW), Familiarity (FAM), Age of Acquisition (AoA), Predictability (PRED), Literality (LIT), Compositionality (COM)

	KNOW	FAM	AoA	PRED	LIT	COM
Adverb Insertion (n = 245)	0.42*	0.29*	-0.18	0.01	-0.16	0.26*
Adjective Insertion (n = 244)	0.28*	0.09	-0.03	-0.02	-0.15	0.18
Left Dislocation (n = 245)	0.41*	0.27*	-0.06	0.06	-0.16	0.29*
Passive (n = 157)	0.25	0.26*	-0.14	-0.06	-0.12	0.19
Movement (n = 241)	0.26*	0.12	-0.01	0.00	-0.02	0.12

* $p < 0.001$

A second purpose of the study was to investigate the correlations among the variables considered. Clearly, familiarity emerges as a crucial factor in the study of idiomatic expressions. It correlates with AoA and influences people's judgments concerning literality, and syntactic flexibility (Libben & Titone, 2008). Other correlations (e.g., between knowledge and compositionality) highlight relations that are still poorly understood and need to be clarified by further research.

Hopefully, the data base will provide a useful tool to all those interested in the representation and processing of idiomatic expressions in Italian.

References

- Ackerman, B. P. (1982). On comprehending idioms: do children get the pictures? *Journal of Experimental Child Psychology*, 33, 439–454.
- Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments and Computers*, 34, 424–434.
- Bates, E., Burani, C., D'Amico, S., & Barca, L. (2001). Word reading and picture naming in Italian. *Memory and Cognition*, 29, 986–999.
- Bianchi, V. (1993). An empirical contribution to the study of idiomatic expressions. *Rivista linguistica*, 5, 349–385.
- Bobrow, S., & Bell, S. (1973). On catching on to idiomatic expressions. *Memory and Cognition*, 1, 343–346.
- Botelho da Silva, T., & Cutler, A. (1993). Ill-formedness and transformability in Portuguese idioms. In C. Cacciari and P. Tabossi (Eds.), *Idioms: Processing, Structure and Interpretation*. Erlbaum.
- Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, 15, 208–216.
- Burani, C., Arduino, L. S., & Barca, L. (2007). Frequency not age of acquisition affects Italian word naming. *The European Journal of Cognitive Psychology*, 19, 828–866.
- Cacciari, C., & Gluckesberg, S. (1991). Understanding idiomatic expressions. The contribution of word meanings. In G. B. Simpson (Ed.), *Understanding Word and Sentence* (pp. 217–240). Amsterdam: North-Holland.
- Cacciari, C., & Levorato, M. C. (1989). How children understanding idioms in discourse. *Journal of Child Language*, 16, 387–405.
- Cacciari, C., Padovani, R., & Corradini, P. (2007). Exploring the relationship between individuals' speed of processing and their comprehension of spoken idioms. *European Journal of Cognitive Psychology*, 25, 668–683.
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27, 668–683.
- Carroll, J. B., & White, M. N. (1973). Word frequency and age-of-acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, 25, 85–95.
- Chomsky, N. (1980). *Rules and representation*. New York: Columbia University Press.
- Colombo, L., & Burani, C. (2002). The influence of age of acquisition, root frequency and context availability in processing nouns and verbs. *Brain and Language*, 81, 398–411.
- Connine, C. M., Blasko, B., Brandt, R., & Kaplan Layer, J. (1992). Idiomatic processing: syntactic frozenness and subjective familiarity. *Psychology Research*, 54, 225–232.
- Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 6, 1084–1096.
- 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.
- Cronk, B. C., & Schweigert, W. A. (1992). The comprehension of idioms: the effects of familiarity, literalness, and usage. *Applied Psycholinguistics*, 13, 131–146.
- Cutting, J. C., & Bock, K. (1997). That's the way of cookie bounces: syntactic and semantic components of experimental elicited idiom blends. *Memory & Cognition*, 25, 57–71.
- Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143–188 (Target article).
- Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1103–1123.
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 515–523.
- Erlich, K., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 20, 641–655.
- Estill, R. B., & Kemper, S. (1982). Interpreting idioms. *Journal of Psycholinguistic Research*, 11, 559–568.
- Fanari, R., Cacciari, C., & Tabossi, P. (2010). The role of idiom length and context in spoken idiom comprehension. *The European Journal of Cognitive Psychology*, 22(3), 321–334.
- Fraser, B. (1970). Idioms within a transformational grammar. *Foundation of Language*, 6, 22–42.
- Frege, G. (1892) 'On concept and object.' Reprinted in B. McGuinness, ed., *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Blackwell, 1984: 182–194.
- Gerhand, S., & Barry, C. (1999). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory and Cognition*, 27, 592–602.
- Gernsbacher, A. M. (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. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory and Cognition*, 8, 449–456.
- Gibbs, R. W. (1987). Linguistic factors in children's understanding of idioms. *Journal of Child Language*, 14, 569–586.
- Gibbs, R. W. (1991). Semantic analysability in children's understanding of idioms. *Journal of Speech and Hearing Research*, 34, 613–620.
- Gibbs, R. W. (1992). What do idioms really mean? *Journal of Memory and Language*, 31, 485–506.
- Gibbs, R. W., & Gonzales, G. P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition*, 20, 243–259.
- Gibbs, R. W., & Nayak, N. P. (1989). Psycholinguistic studies on the syntactic behaviour of idioms. *Cognitive Psychology*, 21, 100–138.
- Gibbs, R. W., Nayak, N. P., Bolton, J. L., & Keppel, M. E. (1989). Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition*, 17, 58–68.
- Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989). How to kick the bucket and not decompose: analyzability and idiom processing. *Journal of Memory and Language*, 28, 576–593.
- Gilhooly, K. J., & Gilhooly, M. L. (1980). The validity of age-of-acquisition ratings. *British Journal of Psychology*, 71, 105–110.

- Glucksberg, S. (1993). Idiom meaning and allusional content. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 3–26). Hillsdale: Erlbaum.
- Hillert, D. (2004). Spared access to idiomatic and literal meanings: a single-case approach. *Brain Language*, 89, 207–215.
- Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E. van den Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and Psychological Perspectives* (pp. 133–166). Hillsdale: Erlbaum.
- Katz, J. J. (1973). Compositionally, idiomacity and lexical substitution. In S. R. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 357–376). New York: Holt, Rinehart & Winston.
- Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalisation from idiom production. *Cognitive Psychology*, 58, 68–101.
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Levorato, M. C., & Cacciari, C. (1992). Children's comprehension and production of idioms: the role of context and familiarity. *Journal of Child Language*, 19, 415–433.
- Levorato, M. C., & Cacciari, C. (1995). The effects of different task on the comprehension and production of idioms in children. *Journal of Experimental Child Psychology*, 60, 261–283.
- Levorato, M. C., Nesi, B., & Cacciari, C. (2004). Reading comprehension and understanding idiomatic expressions: A developmental study. *Brain & Language*, 91, 303–314.
- Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, 36(6), 1103–1121.
- Lyons, A., Teer, P., & Rubenstein, H. (1978). Age-at-acquisition and word recognition. *Journal of Psycholinguistic Research*, 7, 179–187.
- McGlone, S. M., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse Processes*, 17, 167–190.
- Moore, V., & Valentine, T. (1998). The effects of age of acquisition on speed and accuracy of naming famous faces. *The Quarterly Journal of Experimental Psychology*, 51, 485–513.
- Morrison, C. M., & Ellis, A. W. (1995). The role of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 116–133.
- Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology*, 19(2), 167–180.
- Morrison, C. M., Ellis, A. W., & Chappell, T. D. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 50A, 528–559.
- Mueller, R., & Gibbs, R. W. (1987). Processing idioms with multiple meanings. *Journal of Psycholinguistic Research*, 16, 63–81.
- Nippold, M. A., & Martin, S. T. (1989). Idiom interpretation in isolation versus context. A developmental study with adolescents. *Journal of Speech and Hearing Research*, 32, 59–66.
- Nunberg, G. (1978). *The pragmatics of reference*. Bloomington: Indiana University Linguistic Club.
- Nunberg, G., Sag, I., & Wasow, T. (1994). Idioms. *Language*, 70, 491–534.
- Ortony, A., Schallert, D. L., Reynolds, R. E., & Anton, S. J. (1978). Interpreting metaphor and idioms: some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behaviour*, 17, 465–477.
- Papagno, C., Curti, R., Rizzo, S., Crippa, F., & Colombo, M. R. (2006). Is the right hemisphere involved in idiom comprehension? A neuropsychological study. *Neuropsychology*, 20, 598–606.
- Papagno, C., Tabossi, P., Colombo, M. R., & Zampetti, P. (2004). Idiom comprehension in aphasic patients. *Brain and Language*, 89, 226–234.
- Popiel, S. J., & McRae, K. (1988). The figurative and literal senses of idioms, or, all idioms are not used equally. *Journal of Psycholinguistic Research*, 17, 475–487.
- Quartu, B. M. (1993). *Dizionario dei modi di dire della lingua italiana*. Milano: RCS Libri S.p.A.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs: Prentice-Hall.
- Renzi, L., Salvi, G., & Cardinaletti, A. (1988). *Grande grammatica italiana di consultazione: Vol. I. La frase: I sintagmi nominale e preposizionale. [Great Italian Consultation Grammar: Vol. I. The nominal and prepositional phrases]*. Bologna: Il Mulino.
- Schwanenflugel, P. J., & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*, 24, 232–252.
- Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research*, 15, 33–45.
- Schweigert, W. A. (1991). The muddy waters of idiom comprehension. *Journal of Psycholinguistic Research*, 20, 305–314.
- Searle, J. (1975). Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics of speech acts* (pp. 59–82). New York: Academic Press.
- Sidtis Van Lacker, D. (2006). Where in the brain is nonliteral language? *Metaphor & Symbol*, 21, 213–244.
- Sprenger, S., 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, 522–534.
- Tabossi, P., Fanari, R., & Wolf, K. (2005). Spoken idiom recognition: meaning retrieval and word expectancy. *Journal of Psycholinguistic Research*, 34, 465–495.
- Tabossi, P., Fanari, R., & Wolf, K. (2008). Processing idiomatic expressions: effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34 (2), 313–327.
- Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition*, 37(4), 529–540.
- Tabossi, P., & Zardon, F. (1993). The activation of idiomatic meaning during spoken language comprehension. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 145–163). Hillsdale: Erlbaum.
- Tabossi, P., & Zardon, F. (1995). The activation of idiomatic meaning. In M. Everaert, E. van den Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological* (pp. 273–282). New Jersey: Lawrence Erlbaum Associates.
- Tabossi, P., Wolf, K., & Koterle, S. (2009). 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 literality. *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 literality. *Metaphor and Symbolic Activity*, 9(4), 247–270.
- Titone, D. A., & Connine, C. M. (1999). On the compositional and compositional nature of idiomatic expressions. *Journal of Pragmatics*, 31, 1655–1674.
- Turner, J. E., Valentine, T., & Ellis, A. W. (1998). Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision. *Memory and Cognition*, 26, 1282–1291.
- Zola, D. (1984). Redundancy and word perception during reading. *Perception and Psycholinguistics*, 36, 277–284.