Second language learner intuitions of idiom properties: What do they tell us about L2 idiom knowledge and acquisition?

Ferdy Hubers a,b,*, Catia Cucchiariini a, Helmer Strik a,b

a Centre for Language Studies, Radboud University, Nijmegen, The Netherlands
b Centre for Language and Speech Technology, Radboud University, Nijmegen, The Netherlands

Received 18 June 2020; received in revised form 22 July 2020; accepted 23 July 2020
Available online 19 September 2020

Abstract

The present study investigated intuitions of L2 learners about important properties of L2 idiomatic expressions to gain insights for research on L2 idiom processing and acquisition. More specifically, we examined (a) how reliable L2 learners’ intuitions are, (b) how much they differ from native speakers’ intuitions, and (c) whether they are better predictors of L2 idiom knowledge than native speaker intuitions. To this end, Dutch native speakers and German L2 learners of Dutch rated 110 Dutch idioms on frequency of exposure, frequency of use, meaning familiarity, imageability, and transparency and were tested on idiom knowledge. This study shows that L2 learner intuitions about idiom properties are a valuable and reliable source of information on L2 idiom knowledge.

© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Keywords: Idiomatic expressions; L2 learners; Native speakers; Intuitions; Reliability; Idiom knowledge

1. Introduction

Idiomatic expressions like add fuel to the fire and hit two birds with one stone – usually defined as recurrent sequences of words that convey a figurative meaning (Abel, 2003; Cacciari and Glucksberg, 1991; Titone et al., 2015) – appear to be particularly challenging for second language (L2) learners (Cieślicka, 2006; Conklin and Schmitt, 2008; Ellis et al., 2008; Wray, 2000). While such expressions, and formulaic language in general, are pervasive in native language, they are much less frequently used by L2 learners in their L2 (Güngör and Uysal, 2020; Kecskes, 2007; Pawley and Syder, 1983). Even highly proficient L2 learners experience difficulties understanding and using idiomatic expressions (Ellis et al., 2008).

Research has shown that L1 and L2 idiom processing is affected by specific properties of idioms such as frequency, familiarity, transparency, imageability, and L1-L2 similarity (García et al., 2015; Steinel et al., 2007). Data on these idiom properties are usually obtained by collecting people’s intuitions through subjective judgment scales (Bonin et al., 2013; Libben and Titone, 2008; Nordmann et al., 2014; Nordmann and Jambazova, 2017). These can be intuitions subjects have even though they do not know the meaning of an idiom or the intuitions they have formed after learning the meaning of an idiom.

* Corresponding author at: P.O. Box 9103, 6500 HD Nijmegen, The Netherlands.
E-mail address: f.hubers@let.ru.nl (F. Hubers).

https://doi.org/10.1016/j.lingua.2020.102940
0024-3841/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
In general, such intuitions are collected from native speakers of the language under study, who are considered to be the benchmark. When L1 idiom processing is investigated, this is an obvious choice. However, in research on L2 idiom processing and idiom learning, it is not immediately clear from which subjects such intuitions should be collected. On the one hand, it seems reasonable to take the native intuitions as the benchmark, because native speakers can be considered as the model the L2 learners are trying to achieve. On the other, it is conceivable that native intuitions do not reflect L2 knowledge, but rather provide a distorted picture that is not in line with the intuitions and impressions of L2 learners. Since this could have a biasing effect on the results of research, collecting intuitions directly from L2 learners would then seem to be preferable.

Considering that idioms are difficult for L2 learners to acquire and that L2 learners have less experience with the language than native speakers, one might wonder whether L2 learners are at all capable of developing reliable intuitions about idiom properties and whether their intuitions are more informative about their knowledge of idioms than native intuitions. Investigating whether these intuitions are reliable and whether they are more informative about L2 idiom knowledge than native intuitions could contribute to improving the methodology in L2 idiom research.

So far, relatively few studies have investigated intuitions of idiom properties by L2 learners, and generally for a limited number of idiom properties and with little attention to their reliability. Reliability is considered to be the extent to which raters covary or give relative values which are correlated (Rietveld and van Hout, 1993, p. 188). If L2 learners turn out to disagree with each other such that their intuitions are inconsistent, the obtained ratings are not reliable and cannot be used in subsequent analyses. Therefore, reliability is an important aspect of subjective ratings that has to be taken into account.

One study that examined reliability reported low reliability scores for both native and non-native intuitions (Nordmann et al., 2014). Titone and Connine (1994) did not explicitly investigate reliability, but their findings do cast doubt on the reliability of subjective ratings of idiom properties by native speakers. Investigating whether L2 learners are at all capable of developing reliable intuitions about idiom properties and whether these intuitions are related to L2 idiom knowledge would provide an important contribution to L2 idiom research.

Considering the importance of L2 learner intuitions about idiom properties for research on L2 idiom processing, and the scarcity of research on this topic, we conducted a comprehensive study of L2 learner intuitions of idiom properties to investigate (1) how reliable L2 learners’ intuitions about idiom properties are, (2) how L2 intuitions compare to L1 intuitions about idiom properties, and (3) whether L2 intuitions better reflect L2 idiom knowledge than L1 intuitions.

The paper is organized as follows. First, we discuss some important idiom properties and studies that examined the reliability of subjective ratings of idiom properties. Subsequently, we introduce the current study and go on to calculate the reliability of the L1 and L2 intuitions, compare these, and examine to what extent reliable intuitions can be employed to explain L2 idiom knowledge. Finally, we present our results and discuss them in relation to those of previous research.

1.1. Idiom properties

1.1.1. Frequency and familiarity

Idiom frequency is often defined as the frequency with which a speaker or listener indicates to have encountered an idiomatic expression (Carrol et al., 2017; Gembsbacher, 1984; Libben and Titone, 2008), while idiom familiarity is the extent to which people indicate to be familiar with (the meaning of) the idiomatic expression (Abel, 2003; Hubers et al., 2019; Nordmann et al., 2014). Frequency can also be measured objectively from corpora, but this is challenging because of the flexible nature of idiomatic expressions (i.e. different possible word orders and inflections). Only a few studies have compared subjective and objective frequency of units larger than single words (Hubers et al., 2019 for idioms and Siyanova-Chanturia and Spina, 2015 for collocations) and compared these to objective frequencies from corpora. Siyanova-Chanturia and Spina (2015) underline the importance of studying language users’ intuitions about the frequency of units that transcend single words, like collocations and other forms of multiword expressions, as evidence accumulates that these are an important component of language, while still little is known about how they are processed by L2 learners.

Given that individual idioms are not particularly frequent and that, consequently, L2 learners are not likely to encounter them often in naturalistic L2 input (Ellis, 2012), an important question is whether L2 learners have enough opportunities for developing intuitions about idiom frequency and familiarity. Over and above the reduced L2 input, an additional factor that might hinder L2 learners developing intuitions about idioms may be their difficulties in noticing formulaic language. L2 learners are more likely to fail to notice formulaic expressions even when they encounter them (Boers and Lindstromberg, 2012; Eyckmans et al., 2007; Peters, 2012). Idiomatic expressions containing familiar words more often go unnoticed by L2 learners than idiomatic expressions containing unfamiliar words (Kim, 2016; Laufer, 1997).
1.1.2. Transparency

Transparency is generally defined as the degree to which the semantic value of the entire expression can be understood in terms of the semantic values of its constituting words (e.g., Steinel et al., 2007) and is often measured by asking native speakers to indicate to what extent they "consider an idiomatic expression as related to its figurative meaning" (Skoufaki, 2008, p. 20). The idiom spill the beans is opaque, because the figurative meaning to reveal a secret cannot be extracted from the literal interpretation. The expression to hit two birds with one stone is transparent, because the figurative meaning (to solve two problems at once by a single action) can be extracted from the literal interpretation. Transparent idioms appear to pose fewer problems to L2 learners than opaque ones in terms of idiom production and comprehension (Irujo, 1986a; Skoufaki, 2008; Steinel et al., 2007; Yorio, 1989).

Because idiomatic expressions are imbued with specific linguistic and cultural knowledge (Boers et al., 2004; Kövecses and Szabó, 1996), it is to be expected that L1 and L2 transparency intuitions are different. Boers and Webb (2015) compared transparency intuitions of English idioms by native speakers with those of advanced learners of English, and found that the L1 and L2 intuitions were quite different. Abel (2003) investigated intuitions of semantic decomposability (Nunberg, 1978), a concept related to transparency, by L2 learners of English, and concluded that L2 learners tend to rely more on literal meanings than the native speakers in a comparable study of Titone and Connine (1994). Semantic decomposability was assessed by presenting subjects with the idioms and the paraphrases of their figurative meaning and by asking them to rate the idioms as decomposable or non-decomposable (Titone and Connine, 1994). Using the same procedure Nordmann et al. (2014) found that fluent L2 learners of English who had passed the IELTS (International English Language Testing System) proficiency exam judged idioms to be less literal and less decomposable than native speakers. Carrol et al. (2017), on the other hand, reported that native speakers judged English idioms to be less transparent than non-native speakers who had been studying English for an average of 16.7 years and had been living in the UK for an average of 12.6 years. In this respect it is important to notice that Carrol et al. (2017) adopted a different operationalization of transparency and decomposability based on "the stage at which the judgment is being made". In their study, transparency was operationalized as how easily subjects thought they could guess the meaning of the idiom based on the individual words, but without being shown the meaning. Decomposability was defined in the same way, but ratings were obtained later and by showing subjects the correct meaning of the idioms. In between these two questions subjects answered multiple choice items aimed at testing their knowledge of meaning. While these answers gave the authors information about whether the subjects knew the meanings of the idioms, it is still unclear what the subjects were actually judging when they were asked to rate transparency. Because the actual meaning was not shown, they might have had different meanings in mind than the correct one, even a meaning that was not included in the multiple choice items. This complicates the interpretation of the results and the comparison with other studies.

Researchers in cognitive linguistics maintain that transparency intuitions are, at least partly, influenced by inherent properties, like conceptual metaphors and encyclopedic knowledge (Skoufaki, 2008). Keysar and Bly (1995) argued that transparency intuitions are not necessarily rooted in the motivation underlying idioms, but emerge because language users develop explanations for the meanings they have learned to associate with specific idioms. A similar conclusion was drawn by Malt and Elter (2004) with respect to L2 learners. However, Skoufaki (2008) challenged Keysar and Bly’s view, and ascribed their findings in part to specific features of their experiment, over-representation of opaque idioms in their material and a task that pre-empted the use of idiom-inherent properties (Skoufaki, 2008, p. 22). In order to gain more insight into the source of transparency intuitions, Skoufaki (2008) presented advanced L2 learners of English with unknown idiomatic expressions, varying along the transparency dimension, and asked them to guess the meaning and provide an interpretation. She found that high-transparency idioms received fewer different interpretations than low-transparency idioms, which led her to propose a hybrid view of idiom transparency, in which not only idiom familiarity or knowledge affect transparency intuitions, but also idiom-inherent features, i.e. the individual words. A study by Ramonda (2019) on semantic transparency intuitions of idioms by English native speakers also appeared to contradict the highly arbitrary nature of semantic transparency suggested by Keysar and Bly (1995).

The present study systematically compares L1 and L2 transparency intuitions, and investigates these issues making it possible to test different hypotheses. If it is essentially idiom familiarity that drives transparency intuitions, as Keysar and Bly (1995) suggest, then transparency ratings by native speakers should be higher than those by L2 learners. If, on the other hand, transparency intuitions are also affected by intrinsic idiom properties, as proposed by Skoufaki (2008), then it is possible that L2 learners judge the same idioms to be at least as equally transparent as native speakers do. In other words, similar or higher L2 transparency ratings would suggest that transparency intuitions also have a more objective, idiom-inherent basis and are not only induced by idiom familiarity.

1.1.3. Imageability

Imageability indicates the degree to which an idiom can evoke an image (Cacciari and Glucksberg, 1995; Steinel et al., 2007). Cacciari and Glucksberg (1995) found that in native speakers mental images are usually associated with the literal
meaning of idioms rather than with the figurative one. This could imply that the degree to which an image can be formed of an idiom may hamper processing rather than facilitate it.

Research on L2 idiom acquisition has shown that the extent to which idioms can be associated with images has a positive effect on learning the meaning of L2 idioms (SteinEL et al., 2007). This is in line with the dual coding hypothesis (Pavio, 1988; Sadoski, 2005), which assumes that cognition occurs in a verbal code for language and a non-verbal code for mental imagery. However, Boers et al. (2008) found that pictorial elucidation was not conducive to better retention of the linguistic form of the idioms.

The present study makes a direct comparison between L1 and L2 imageability intuitions. In addition, by investigating the impact of imageability on idiom knowledge we expect to gain a better understanding of the processes underlying L2 idiom acquisition.

1.1.4. Cross-language overlap

Studies on L2 idioms also need to take account of the extent to which L2 idioms exist in L1 (Charteris-Black, 2002; CieślIcka, 2006, 2015; Irujo, 1986a,b; Liontas, 2002, 2015). Previous research indicates that idioms that exist in both L1 and L2 are less difficult (Carrol et al., 2017; Irujo, 1986a; Laufer, 2000), but this seems to hold especially for comprehension, while in production similar idioms may be more problematic because of interference from L1 (Charteris-Black, 2002; Liontas, 2002; Yorio, 1989).

The extent to which L2 idioms exist in the L1 can also influence L2 intuitions about idiom properties. L2 idioms with exact equivalents in the L1 appear to be judged as more familiar and more transparent than L2 idioms which do not have identical matches in the L1 (Carrol et al., 2017). It is not clear, however, how L2 intuitions are influenced by intermediate levels of cross-language overlap and how cross-language overlap affects the relation between subjective and objective characteristics of L2 idioms.

A more detailed classification that takes account of both form and meaning as proposed by Titone et al. (2015) seems to be required to obtain a clearer understanding of how cross-language overlap affects L2 intuitions of idiom properties. These authors used a scale ranging from 1 to 5 and found that cross-language overlap facilitated idiom processing. The current study examines cross-language overlap and relates this to L2 idiom knowledge.

1.2. Reliability

L2 learners are generally less exposed to the L2 and in particular to L2 idioms (Wray, 2002) than native speakers. As a result, they are likely to develop less reliable intuitions about idiom familiarity, frequency, transparency, and imageability. However, this might be modulated by the proficiency level of the L2 learners, the amount of L2 experience, and their native language.

Nordmann et al. (2014) investigated L1 and L2 intuitions and examined their reliability by collecting ratings of familiarity, meaning, literality, and decomposability through 7-point Likert scales from 44 native speakers and 32 non-native speakers of English for 100 English idioms. The authors analyzed the reliability of the ratings and concluded that it was low for both L1 and L2 intuitions. The diversity among the non-native speakers’ native languages might have caused differences in the ratings that affected the degree of reliability. In a more homogeneous sample of participants with the same L1 reliability should be higher, although this might seem less plausible given the low reliability values that Nordmann et al. (2014) obtained for native speakers, who constitute a more homogeneous group. The present study will throw more light on this issue.

1.3. The present study

The review of previous research on L2 idiom learning and the role of L2 intuitions about idiom properties reveals that a number of important and crucial questions remain unanswered. These concern the reliability of L2 intuitions, the differences between L1 and L2 intuitions, and their possible consequences for subsequent research on L2 idiom processing. Moreover, it is not yet clear how L1 and L2 intuitions and cross-language overlap are related to an objective measure of L2 idiom knowledge.

To investigate these issues, we collected intuitions of frequency, familiarity, usage, transparency, and imageability of Dutch idiomatic expressions from German learners of Dutch and Dutch native speakers, data on objective frequency of idioms from corpora and on objectively assessed idiom meaning recognition as a measure of idiom knowledge. Using a test of idiom knowledge provides more direct measurements than those obtained in previous approaches in which idiom knowledge was estimated based on self-reported data and/or familiarity judgments, or data collected from other, comparable subjects (Beck and Weber, 2016; CieślIcka, 2013; Nordmann et al., 2014; Titone and Connine, 1994). This test-based method, which was proposed in van Ginkel et al. (2016), and Hubers et al. (2016) for Dutch idioms was later
adopted by Carrol et al. (2017) in their study on native and non-native understanding of figurative phrases from English, German, Bulgarian and Chinese. In the remainder of this paper we refer to this specific type of idiom knowledge as L2 idiom knowledge.

We addressed the following research questions:

1. Are L2 learners capable of developing reliable intuitions about idiom properties?
2. How do L2 intuitions compare to L1 intuitions about idiom properties?
3. Do L2 intuitions better reflect L2 idiom knowledge than L1 intuitions?

As to RQ1, we hypothesize that L2 learners are capable of developing reliable intuitions about frequency, usage, familiarity, imageability, and transparency, but that these are less reliable than L1 intuitions, since L2 learners are much less exposed to the target language and culture. Although Nordmann et al. (2014) found low reliability of L2 intuitions of idiom properties, we hope to increase the chance of obtaining reliable results by adopting a more suitable statistical measure of reliability, more specific questions about the idiom properties under study, since this can influence the results (Hubers et al., 2019), and a relatively homogenous sample of L2 learners with the same L1.

For RQ2, we expect that limited L2 exposure leads to lower ratings for familiarity, frequency, and usage by L2 learners than by native speakers. As to transparency, we are interested in comparing the predictions by Keysar and Bly (1995) with those made by Skoufaki (2008). If transparency is mainly influenced by idiom familiarity, as Keysar and Bly (1995) suggest, then we expect higher transparency ratings for native speakers as compared to L2 learners. However, if transparency intuitions are also affected by idiom intrinsic properties, as proposed by Skoufaki (2008), then native speakers are expected not to judge idioms as more transparent than L2 learners do. For imageability, the picture is less clear-cut because research is limited. Native speakers tend to associate mental images with the literal meanings of idioms rather than with the figurative ones (Cacciari and Glucksberg, 1995). Given their higher proficiency, they should be more likely to link idioms to images. The limited research on the role of imageability in idiom learning shows a facilitative role suggesting that L2 learners might exploit this more than native speakers do, because subjects who “found it easy to conjure up an image during learning performed well during testing” according to Steinel et al. (2007, p. 479).

To gain more insight into the relationship between the L1 and L2 intuitions, we will also check the correlations between these ratings. Intuitions of frequency, familiarity, and usage are more experience-based while intuitions of transparency and imageability are more related to intrinsic properties of the idioms themselves (Hubers et al., 2019). For these reasons we should expect stronger correlations between L1 and L2 intuitions of transparency and imageability, than for those of frequency, familiarity, and usage.

With respect to RQ3, our hypothesis is that if L2 intuitions turn out to be sufficiently reliable, they are also better predictors of L2 idiom knowledge than L1 intuitions. Furthermore, on the basis of the findings from Titone et al. (2015), we expect a positive effect of cross-language overlap on L2 idiom knowledge.

2. Method

2.1. Participants

Native speakers. 26 Dutch native speakers participated in our study (24 females). They were mainly university students, were on average 22.7 years old, ranging from 19 to 34 (SD = 3.2).

L2 learners. 26 German learners of Dutch participated in our study (23 females). They studied or worked at a Dutch university, were between 21 and 32 years old (mean age = 24.76, SD = 3.46), had started learning Dutch around the age 18 to 20 (see Table 1), and reported to be moderately to highly proficient in Dutch. To further check their proficiency, we administered LexTale, an efficient and valid test of vocabulary knowledge ranging from 0 to 100 (Lemhöfer and Broersma, 2012). Their average score of 69.04 (SD = 11.75) confirmed that they were moderately to highly proficient in Dutch. By way of comparison, the native speakers obtained an average score of 90.82 (SD = 6.07) on the same test.

This study was ethically assessed and approved of by the Ethics Assessment Committee (EAC) of the Faculty of Arts and the Faculty of Philosophy, Theology and Religious Studies of Radboud University Nijmegen (number 3382).

2.2. Materials

We selected 110 Dutch idioms from a database consisting of 393 idiomatic expressions rated by native speakers on various idiom properties, such as Familiarity, Transparency, and Imageability (Hubers et al., 2018). Idioms were selected to obtain a plausible reflection of the variation in the full dataset. To design multiple choice items for the knowledge test we created three incorrect alternative meanings that would be plausible if one were not familiar with the idiom.
Cross-language overlap. For the 110 Dutch idioms, the degree of similarity between Dutch and German was determined by two bilingual German-Dutch students. They assessed cross-language overlap using a slightly adapted version of the rating system described in Titone et al. (2015). Four levels of overlap were distinguished: (1) The Dutch idiom has no equivalent in German (NE), (2) The Dutch idiom has an equivalent in German, but in completely different content words (DW), (3) The Dutch idiom has an equivalent in German that has n content words in common (nW), (4) The Dutch idiom has a word-to-word equivalent in German (AW). The students individually scored all idioms and subsequently compared their scores.

Objective idiom frequency. Objective idiom frequencies were obtained from the 500-million-words SoNaR corpus (Oostdijk et al., 2013), a corpus of written Dutch. First, we identified one content word per idiom (usually a noun) and extracted all sentences from the corpus containing this content word. For example, we looked for all sentences containing the Dutch word lamp “lamp” in the corpus (from the Dutch idiom tegen de lamp lopen “to get caught”). Second, we obtained the sentences containing the idiomatic expressions in the subset by means of pattern matching, taking into account different word orders and inflections of the verb.

2.3. Design and procedure

Operationalization of the idiom properties. Five subjective idiom properties were measured through 5-point Likert scales and objective knowledge of idiom meaning through a multiple choice test.

Frequency: relative degree to which participants indicate to have come across an idiom in speech or in print (Titone and Connine, 1994). Usage: frequency with which subjects indicate having used an idiom. Familiarity: how well speakers say that they know the meaning of an idiom (Nordmann et al., 2014: 88). Imageability: the extent to which an idiom can evoke an image (Boers et al., 2008; Steinel et al., 2007). This image could be based on the literal or the figurative meaning. Transparency: the extent to which the original metaphorical motivation of an idiomatic phrase can be deduced from its literal analysis (Cieślacka, 2015, p. 213). Objective idiom knowledge: obtained from a multiple-choice test of meaning recognition.

Questionnaire. The rating study was conducted online through the Qualtrics platform (Qualtrics, 2005). The participants started by filling in a background questionnaire and completed the Dutch version of the LexTale vocabulary test (Lemhöfer and Broersma, 2012), as an indicator of their proficiency in Dutch.

In the main part of the rating study the participants had to answer five questions about the idiomatic expressions on 5-point Likert scales (questions 1, 2, 3, 4, and 6), and one multiple choice item (question 5):

1. Frequency: How often have you heard or read this expression?
   - (1. very rarely–5. very often)
2. Usage: How often have you used this expression yourself?
   - (1. very rarely–5. very often)
3. Familiarity: How familiar are you with the meaning of this expression?
   - (1. completely unfamiliar–5. completely familiar)
4. Imageability: How easily can you form an image of this expression?
   - (1. very hard–5. very easily)
5. Objective idiom knowledge (recognition): Which definition is the correct one?
   - (multiple choice question: 4 alternatives)
6. Transparency: How clear is the meaning of this expression based on the individual words in the expression?
   - (1. very unclear–5. very clear)
In a within-subject design, ratings on one dimension may be influenced by ratings on the other dimensions. However, Nordmann and Jambazova (2017) found no effects of study design (within-subjects vs. between-subjects) on idiom ratings. Moreover, “it is important to collect these ratings within subjects, because they can never be independent and should not be treated as such” (Nordmann and Jambazova, 2017, p. 200).

The idioms were organized in four blocks of 27, 28, 28, and 27 expressions respectively and the order of presentation within blocks was randomized. It took the participants between 30 and 45 min to complete each block.

2.4. Data analysis

To address RQ1, we examined the reliability of the L1 and L2 intuitions by calculating the Intraclass Correlation Coefficient (ICC) using the ‘rel’ package (Lo Martire, 2017) in R, version 3.4.0 (R Development Core Team, 2008). The ICC was calculated for the averaged ratings with the parameters 'two-way', and 'absolute agreement', (random effects for participants and items). We also examined the reliability of the objective idiom knowledge test by calculating Cronbach's alpha using the same R package (parameters 'two-way' and 'consistency').

To answer RQ2 we then compared the L1 and L2 intuitions by computing the mean ratings and standard deviations for all subjective dimensions for native speakers and L2 learners separately. The proportions correct on the multiple choice question were taken to calculate the average objective idiom knowledge and standard deviation. An independent samples T-test was performed to assess the differences between the groups. Pearson's correlations were calculated between the L1 and L2 intuitions, and between the L1 and L2 idiom knowledge based on aggregated data.

To examine to what extent L1 or L2 intuitions better reflect L2 idiom knowledge (research question 3), we performed logistic mixed effects regression analyses based on the individual data. We used the statistical software package ‘R’, version 3.4.0 (R Development Core Team, 2008), and the R packages ‘lme4’ (Bates et al., 2015), ‘lmerTest’ (Kuznetsova et al., 2017), and ‘effects’ (Fox, 2003) to conduct the analyses. The models were built in a forward manner. Since L1 intuitions are often taken as benchmarks in idiom processing studies, we started off with an initial model including a random intercept for participants and fixed effects of the idiom properties under study as rated by the native speakers. Subsequently, we added the ratings based on L2 intuitions (as fixed factors) one by one to the model and examined whether the model fit improved. If this was not the case, the predictor was not included in the model. Next, objective frequency, cross-language overlap, participant characteristics (fixed factors) and potential random factors were added using the same procedure. If the model fit did not improve, the predictor was not included in the model. During this process, we also excluded predictors that did not significantly contribute to the model fit. Both the initial and the final model, based on L1 and L2 intuitions, respectively, are reported in this paper.

3. Results

3.1. Reliability

Table 2 shows the ICC for the various idiom properties included in our rating study per participant group. Cronbach's alpha is presented as a reliability measure of the idiom knowledge test. The L1 and L2 ratings for all dimensions were highly reliable (ICC > 0.9, and ICC > 0.85 respectively), as well as the L1 and L2 performance on the objective idiom knowledge test (Cronbach's alpha = 0.91 for both groups).

3.2. Comparison of L1 and L2 intuitions

Table 3 presents the mean ratings of the various dimensions and their standard deviations as provided by the Dutch native speakers and the German learners of Dutch L2. An independent t-test showed significant differences between the
Table 3
Mean and SD for ratings on idiom properties and for performance on knowledge question for native speakers and L2 learners.

| Idiom property       | Native speakers | L2 learners | t value (df) | d     |
|----------------------|-----------------|------------|-------------|-------|
| Frequency            | 3.77 (0.80)     | 1.75 (0.61) | 21.077 (203.3)** | 2.84  |
| Familiarity          | 3.43 (0.84)     | 1.91 (0.78) | 13.906 (218)** | 1.88  |
| Usage                | 2.41 (0.87)     | 1.26 (0.34) | 12.992 (142.0)** | 1.75  |
| Transparency         | 2.54 (0.76)     | 3.03 (0.80) | -4.720 (218)** | 0.64  |
| Imageability         | 3.29 (0.80)     | 2.53 (0.73) | 7.344 (218)** | 0.99  |
| Objective idiom knowledge | 0.88 (0.18)     | 0.62 (0.26) | 8.877 (192.7)** | 1.17  |

** p < .001, n = 110.

Table 4
Pearson’s correlations between L1 and L2 ratings per idiom property.

|                      | L1–L2 |
|----------------------|-------|
| Familiarity          | 0.195*|
| Frequency            | 0.254**|
| Usage                | 0.356***|
| Transparency         | 0.649***|
| Imageability         | 0.586***|
| Objective idiom knowledge | 0.563***|

*p< .05, **p< .01, ***p< .001, n= 110.

L1 and L2 ratings on all dimensions. The most pronounced differences were observed in the frequency, familiarity, and usage dimensions, which were assigned much lower values by the L2 learners, and as witnessed by the very large effect sizes (Cohen’s $d > 1.5$; Sawilowsky, 2009). In addition, the L2 learners’ knowledge was much lower than that of the native speakers.

We examined the Pearson’s correlations between the L1 and L2 intuitions for the different idiom properties (see Table 4). For all dimensions, significant correlations were observed. L1 and L2 transparency intuitions showed the strongest correlation (Pearson’s $r = .65$). High correlations were also observed for imageability (Pearson’s $r = .59$) and objective idiom knowledge (Pearson’s $r = .56$), while much lower correlations were found for the dimensions familiarity (Pearson’s $r = .20$), frequency (Pearson’s $r = .25$), and usage (Pearson’s $r = .36$).

3.3. Intuitions and objective idiom knowledge

To examine to what extent L1 and L2 intuitions of idiom properties reflect L2 idiom knowledge, we carried out logistic mixed effects regression analyses. The responses to the multiple-choice question on idiom knowledge by the L2 learners were converted into a binary variable expressing whether the multiple-choice question was answered correctly or not. This binary variable was used as the dependent variable in the regression analyses.

As explained above, we started off with an initial model only including native predictors: (1) L1 Familiarity, (2) L1 Transparency, (3) L1 Imageability. In addition, we included Participants (random intercept only), and Idioms (random intercept only) as a random effect. See Table 5 for the final model. Because of multicollinearity the predictors L1 Usage and L1 Frequency could not be included in the model.

Transparency as judged by native speakers turned out to be a significant predictor of L2 idiom knowledge. After having established the initial model, we added the same dimensions as rated by the L2 learners. In the presence of these dimensions, the L1 intuitions no longer significantly contributed to the model, and were therefore removed.

The final model included the following predictors: (1) L2 Familiarity, (2) L2 Transparency, (3) L2 Imageability, (4) LexTale score, and (5) Cross-language overlap (reference category: AW). We also added Idioms (random intercept only), and Participants as random effects (random intercept and random slopes of L2 Imageability and L2 Transparency) and standardized the predictors (see Table 6).

Because L2 Frequency, L2 Usage and Objective Frequency did not improve the model fit, they were excluded. This was also the case for interactions of intuitions of idiom properties, and language background variables, such as Number of hours speaking Dutch a week, and Number of years living in the Netherlands.
Table 5
Initial regression model with L2 objective idiom knowledge as the dependent variable and L1 variables as predictors.

| Fixed effects          | Beta    | Std. error | z Value |
|------------------------|---------|------------|---------|
| (Intercept)            | 0.7052  | 0.1911     | 3.690***|
| L1 Familiarity         | 0.1390  | 0.1364     | 1.019   |
| L1 Transparency        | 0.7581  | 0.1650     | 4.595***|
| L1 Imageability        | −0.0668 | 0.1605     | −0.416  |

Random effects

|          | Variance | Std. deviation |
|----------|----------|----------------|
| Idiom    | Intercept| 1.6310         | 1.2772   |
| Participant| Intercept| 0.4950         | 0.7035   |

*** p < .001.

Table 6
Final regression model with L2 objective idiom knowledge as the dependent variable and L2 variables as predictors.

| Fixed effects          | Beta    | Std. error | z Value |
|------------------------|---------|------------|---------|
| (Intercept)            | 1.9217  | 0.3383     | 5.681***|
| L2 Familiarity         | 0.3976  | 0.1083     | 3.672***|
| L2 Transparency        | 2.4547  | 0.2034     | 12.069***|
| L2 Imageability        | −0.5693 | 0.1484     | −3.836***|
| LexTale                | 0.6134  | 0.1640     | 3.741***|
| Cross-language overlap NE| −0.8051 | 0.3862     | −2.085*  |
| Cross-language overlap DW| −0.9637 | 0.3620     | −2.662***|
| Cross-language overlap nW| −0.7738 | 0.3428     | −2.257*  |

Random effects

|          | Variance | Std. deviation | Corr.   |
|----------|----------|----------------|---------|
| Idiom    | Intercept| 0.9277         | 0.9632  |
| Participant| Intercept| 0.6155         | 0.7820  |
|          | L2 Transparency| 0.7204 | 0.8488 | 0.17 |
|          | L2 Imageability| 0.2486 | 0.4986 | 0.26 | −0.71 |

*p<.05, **p<.01, ***p<.001.

Positive effects were found for L2 Familiarity ($\beta = 0.40$, $SE = 0.11$, $p < .001$), and L2 Transparency ($\beta = 2.45$, $SE = 0.20$, $p < .001$), while for L2 Imageability a negative effect was found ($\beta = −0.57$, $SE = 0.15$, $p < .001$). We observed a positive effect of vocabulary knowledge as measured by LexTale ($\beta = 0.61$, $SE = 0.16$, $p < .001$) and of cross-language overlap. More specifically, if a Dutch idiom was a word-by-word translation of the German expression (AW), L2 learners more often selected the correct meaning in the multiple-choice question than if the Dutch expression had no German equivalent at all (NE) ($\beta = −0.81$, $SE = 0.39$, $p < .05$), the German equivalent consisted of completely different words (DW) ($\beta = −0.96$, $SE = 0.36$, $p < .01$) or if the German equivalent had a number of words in common with the Dutch idiom, but was not a word-by-word translation (nW) ($\beta = −0.77$, $SE = 0.34$, $p < .05$). However, relevelled versions of the model showed that the categories NE, DW, and nW did not significantly differ from each other (NE-DW: $\beta = −0.16$, $SE = 0.32$, $p = .62$; NE-nW: $\beta = 0.03$, $SE = 0.31$, $p = .92$; DW–nW: $\beta = 0.19$, $SE = 0.28$, $p = .49$).

4. Discussion

4.1. Reliability

L2 intuitions about idiom properties turned out to be highly reliable (ICC > .86) and the reliability coefficients were only slightly lower than those obtained for L1 intuitions (ICC > .91). The objective idiom knowledge test also turned out to be reliable for both native speakers and L2 learners (Cronbach’s alpha = .91).

These findings on the reliability of L1 and L2 intuitions are in contrast with those by Nordmann et al. (2014), who reported low reliability for L1 and L2 intuitions. The L2 learners in their study formed a less homogenous group than those in our study, who all had German as their L1, a language that is relatively close to Dutch. It might be argued that this greater
homogeneity led to the high reliability coefficients, but this does not appear to be the only explanation, as the native speakers in Nordmann et al. (2014) also constituted a homogeneous group and reliability was low also for that group of participants. A possible explanation for this difference might be the measure used to calculate reliability. Nordmann et al. (2014) used Krippendorff’s alpha and interpreted this as a measure of reliability. However, this measure in fact reflects agreement rather than reliability (Tinsley and Brown, 2000). Agreement and reliability measure different aspects of a set of ratings. Agreement has to do with the absolute values of ratings, it indicates to what extent the values are identical. Reliability, on the other hand, indicates to what extent ratings covary (Rietveld and van Hout, 1993; Tinsley and Weiss, 1975). Therefore, low values of Krippendorff’s alpha may indicate low agreement, but not necessarily low reliability (see Hubers et al., 2019 for a more elaborate discussion on reliability).

Another important element that probably contributed to the high reliability of our ratings for the native speakers and the L2 learners was the precise and careful way in which the questions about the idiom properties were formulated. Previous studies used varying definitions of the idiom properties under investigation and the questions posed were sometimes ambiguous, leaving room for different interpretations, which may result in more variation in the ratings and lower reliability (see Hubers et al., 2019).

4.2. Comparison of L1 and L2 Intuitions

L1 intuitions about frequency, familiarity, and usage exhibited much higher values than those of L2 learners. This is in line with our expectations and with previous studies reporting large differences between L1 and L2 ratings of idiom frequency (Carrol et al., 2017; Nordmann et al., 2014) and familiarity (Nordmann et al., 2014). L2 learners have less experience with the second language than native speakers of that language, and thus with idiomatic expressions (Wray, 2002).

Although a number of studies have examined idiom imageability in native speakers (Cacciari and Glucksberg, 1995), and L2 learners (Steinell et al., 2007), to the best of our knowledge a systematic comparison of L1 and L2 ratings on this dimension has not been conducted before. Based on the limited research on the role of imageability in L2 idiom learning, we expected the L2 learners to rely more on images than the native speakers. However, we found that the latter group rated the idioms as more imageable than the L2 learners did. Apparently, the native speakers’ higher language proficiency and higher familiarity with the meaning of the idioms made it easier for them to visualize the idioms.

The L2 transparency ratings were higher than the L1 transparency ratings, which is in contrast with studies reporting that native speakers consider idioms to be more transparent than L2 learners do (Abel, 2003; Malt and Eiter, 2004). Keysar and Bly (1995) argued that transparency intuitions emerge post factum, after participants have learned to associate a given meaning with an idiom, suggesting that transparency intuitions are not necessarily derived from literal meanings. If this is true and transparency intuitions are indeed driven by idiom knowledge, then L1 transparency intuitions should be higher than L2 transparency intuitions. However, our results show the opposite pattern. In our study, we selected the idioms in such a way that transparency, as judged by native speakers in an earlier study, varied. For our question about transparency, we presented the idiom and its meaning and asked people to what extent the individual words could be used to arrive at the figurative meaning. Due to this operationalization people were encouraged to use idiom-inherent properties to rate transparency. Our finding that native speakers did not assign higher transparency ratings to the same idioms than L2 learners therefore lends support to Skoufaki’s (2008) more hybrid view of idiom transparency. She suggests that transparency intuitions are also based on more “objective”, idiom-inherent properties and are not only developed after participants have learned to associate a specific meaning with an idiom. The higher transparency ratings by L2 learners also suggest that the L2 learners focused more on the individual words in the idioms and on their contribution to the meaning of the idiom. This is in line with Abel (2003), who found that L2 learners judged non-transparent idioms as more transparent. This, in combination with the finding that transparency was the most important predictor of L2 idiom knowledge (see 4.3), seems to indicate that L2 learners do indeed rely more on the individual words than native speakers, as was suggested by Cieślicka (2006). Finally, the high reliability coefficients obtained for the transparency ratings suggest that L2 learners are capable of consistently rating these intrinsic idiom properties.

The largest differences between L1 and L2 intuitions are observed for the dimensions frequency, familiarity, and usage, whereas for the dimensions transparency and imageability the differences are much smaller. This dichotomy is also visible in terms of correlations: The L1 and L2 intuitions of the dimensions transparency, and imageability are more strongly correlated with each other than the intuitions of the dimensions familiarity, frequency and usage. This suggests that the dimensions frequency, familiarity, and usage are different from transparency and imageability. This difference may lie in the nature of the dimensions. Intuitions of frequency, familiarity, and usage are based on people’s experience with the idiom (experience-based dimensions), whereas transparency and imageability intuitions are more related to intrinsic properties of the idioms themselves (content-based dimensions) (see Hubers et al., 2019). It is therefore plausible
that the largest differences between the native speakers and L2 learners (in terms of both the mean ratings and correlations) are observed for the dimensions that are related to language experience.

4.3. Intuitions and objective idiom knowledge

Researchers in L2 idiom processing and acquisition often rely on L1 intuitions as a basis for material selection in experiments targeting L2 learners and statistical analyses about L2 idiom processing and learning. We investigated whether L1 intuitions are good predictors of L2 idiom knowledge, or whether L2 intuitions would be preferable. L1 intuitions of familiarity and transparency did affect L2 idiom knowledge in the absence of L2 intuitions. However, after adding the L2 intuitions of the corresponding idiom properties to the analysis, the L1 intuitions lost their predictive power. In other words, L2 intuitions of familiarity, transparency, and imageability seem to be more informative when studying L2 idiom knowledge than L1 intuitions, at least for a relatively homogeneous group of L2 learners with the same L1 as in the present study.

The analyses do not only allow us to examine whether L2 or L1 intuitions are better predictors of L2 idiom knowledge, but also give insights into the nature of the relations between L2 idiom knowledge and the intuitions. The final analysis revealed that familiarity, transparency, and imageability (as rated by the L2 learners) affected L2 idiom knowledge. For familiarity and transparency positive effects were observed, while imageability negatively affected L2 idiom knowledge. Transparency most strongly influenced L2 idiom knowledge. L2 learners rely on idiom transparency to arrive at the idiom meaning, because they are less familiar with the expressions than native speakers.

The more transparent the idiom, the better the L2 idiom knowledge. Since L2 learners are in general less familiar with the meaning of the idioms, they are more likely to visualize the literal reading of the idioms. In turn this could hinder them to correctly answer the knowledge question. The negative effect of imageability might seem to contrast with findings indicating that forming an image of the idiom positively affects idiom learning (Steinell et al., 2007). However, as described in the introduction, negative effects of imageability on idiom processing and idiom knowledge have also been observed for native speakers, who have much more experience with idioms (Cacciari and Glucksberg, 1995; Hubers et al., 2019) and L2 learners (Boers et al., 2008). These studies suggest that people are more inclined to form an image of the literal reading of an idiom.

Vocabulary knowledge (LexTale) and cross-language overlap positively affected L2 idiom knowledge. Although we recruited a homogeneous group of L2 learners in terms of language background and proficiency, vocabulary knowledge was still an important predictor of L2 idiom knowledge. This finding confirms the strong relation between vocabulary knowledge and idiom knowledge, as other studies also pointed out (e.g., Irujo, 1986a; Zyzik, 2011). Zyzik (2011) found that lexical knowledge of single words facilitated idiom learning: Meaning recall for idioms containing unknown words was more difficult than for idioms containing known words.

The effect of cross-language overlap indicates that L2 learners benefit from idioms in their L1 that are word-to-word translations of the idiom in their L2. Surprisingly, we did not find significant differences between the other three categories. Idiomatic expressions that do exist in L1 as an almost, but not exact word-to-word translation did not appear to be better known than idiomatic expressions that do not exist in L1, or do exist in L1, but in totally different words. L2 learners appear to use their L1 idiom knowledge to arrive at the correct idiom meaning in the L2 especially for L2 idioms that have an exact equivalent in the L1. In the situation of the exact equivalents, L2 learners probably feel confident enough to rely on their L1, whereas if the L2 idioms only partially overlap with the L1 equivalents, they are hesitant to resort to their L1 knowledge. These findings complement those of other studies on L2 idioms (Carroll et al., 2017; Charteris-Black, 2002; Irujo, 1986a; Titone et al., 2015) and provide a more fine-grained picture of the impact of cross-language overlap on L2 idiom knowledge.

L2 intuitions of frequency and usage did not significantly affect L2 idiom knowledge. This is in contrast to Hubers et al. (2019), who reported significant effects of L1 frequency and usage on L1 idiom knowledge. The absence of these effects may be due to the relatively low scores and limited variability in the experience-based dimensions familiarity, frequency and usage as rated by the L2 learners, suggesting that familiarity, frequency and usage measure roughly the same construct.

Interestingly, even in the absence of an effect of L2 subjective frequency, we did not find an effect of objective frequency as obtained by corpora. Although an objective measure of idiom frequency positively affects idiom knowledge of native speakers (Hubers et al., 2019), it does not seem a relevant factor in predicting idiom knowledge by L2 learners. Objective idiom frequency may start to positively affect idiom knowledge only after more exposure to the L2. The absence of an effect of objectively measured idiom frequency may be a reason to, in the case of L2 learners, rely more on intuitions obtained from the learners themselves.

5. Conclusions

In the current study, we investigated to what extent L2 intuitions of idiom properties differ from L1 intuitions in terms of average values and reliability, and whether L1 or L2 intuitions are better predictors of L2 idiom knowledge.
The results presented here provide relevant insights that should be taken into account when designing experiments on L2 idiom processing. It is often the case that in such studies the selection of material and the statistical analyses are based on intuitions obtained from L1 speakers. Our study has shown that L1 intuitions are different from L2 intuitions and that the latter are reliable and better reflect L2 knowledge, at least when they are obtained from a relatively homogeneous group of L2 learners with the same L1. Consequently, we recommend that researchers in L2 idiom processing and idiom learning do not take L1 intuitions for granted. Depending on the specific aim of the research, it might be worthwhile to collect and use L2 ratings of idiomatic expressions. In addition, when collecting these ratings, it is important to precisely formulate the rating questions in order to obtain reliable results.

To conclude, the current study shows that L2 intuitions about idiom properties are a valuable and reliable source of information that gives more insight into L2 idiom knowledge. Differences between intuitions by L2 learners and native speakers and their relations to idiom knowledge lead us to conclude that for L2 learners the individual words are more salient than the figurative meaning, whereas for native speakers this does not seem to be the case. These differences and the finding that L2 learners are able to develop reliable intuitions suggest that L2 intuitions deserve more attention and should be taken into account when studying L2 idiom processing and acquisition.

Funding

This work is part of the research program Free Competition in the Humanities with project number 360-70-510 NWO ISLA, which is financed by the Netherlands Organization for Scientific Research (NWO).

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We are grateful to our fellow members of the ISLA team, Wendy van Ginkel, Ton Dijkstra, Louis ten Bosch, and Alexander Wahl for their valuable input and to Ton Dijkstra, Roeland van Hout and Rob Schoonen for their helpful comments on earlier versions of the manuscript.

References

Abel, B., 2003. English idioms in the first language and second language lexicon: a dual representation approach. Second Lang. Res. 19 (4), 329–358. http://dx.doi.org/10.1111/0267658303sr226oa
Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1), 1–48. http://dx.doi.org/10.18637/jss.v067.i01
Beck, S.D., Weber, A., 2016. Bilingual and monolingual idiom processing is cut from the same cloth: the role of the L1 in literal and figurative meaning activation. Front. Psychol. 7, 1–16. http://dx.doi.org/10.3389/fpsyg.2016.01350
Boers, F., Demecheleer, M., Eyckmans, J., 2004. Cross-cultural variation as a variable in comprehending and remembering figurative idioms. Eur. J. Engl. Stud. 8 (3), 375–388. http://dx.doi.org/10.1080/1382557042000277449
Boers, F., Lindstrømberg, S., 2012. Experimental and intervention studies on formulaic sequences in a second language. Annu. Rev. Appl. Linguist. 32, 83–110. http://dx.doi.org/10.1017/S0267190512000050
Boers, F., Lindstrømberg, S., Littlemore, J., Stengers, H., Eyckmans, J., 2008. Variables in the mnemonic effectiveness of pictorial elucidation. In: Boers, F., Lindstrømberg, S. (Eds.), Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology. Mouton De Gruyter, pp. 189–216.
Boers, F., Webb, S., 2015. Gauging the semantic transparency of idioms: do natives and learners see eye to eye? In: Heredia, R.R., Cieslacka, A. B. (Eds.), Bilingual Figurative Language Processing. Cambridge University Press, pp. 368–392.
Bonin, P., Méot, A., Bugaiska, A., 2013. Norms and comprehension times for 305 French idiomatic expressions. Behav. Res. Methods 45, 1259–1271. http://dx.doi.org/10.3758/s13428-013-0331-4
Cacciari, C., Glucksberg, S., 1991. Understanding idiomatic expressions: the contribution of word meanings. In: Simpson, G.B. (Ed.), Understanding Word and Sentence. North-Holland, pp. 217–240. http://dx.doi.org/10.1016/S0166-4115(08)61535-6North-Holland.
Cacciari, C., Glucksberg, S., 1995. Understanding idioms: do visual images reflect figurative meanings? Eur. J. Cognit. Psychol. 7 (3), 283–305. http://dx.doi.org/10.1080/09541449508402450
Carrol, G., Littlemore, J., Dillon Dowens, M., 2017. Of false friends and familiar foes: comparing native and non-native understanding of figurative phrases. Lingua. http://dx.doi.org/10.1016/j.lingua.2017.11.001
Charteris-Black, J., 2002. Second language figurative proficiency: a comparative study of Malay and English. Appl. Linguist. 23, 104–133.
Cieslacka, A.B., 2006. Literal salience in on-line processing of idiomatic expressions by second language learners. Second Lang. Res. 22 (2), 115–144. http://dx.doi.org/10.1111/j.0267658306sr263oa
Cieslacka, A.B., 2013. Do nonnative language speakers chew the fat and spill the beans with different brain hemispheres? Investigating idiom decomposability with the divided visual field paradigm. J. Psycholinguist. Res. 42 (6), 475–503. http://dx.doi.org/10.1007/s10936-012-9232-4
Cieślicka, A.B., 2015. Idiom acquisition and processing by second/foreign language learners. In: Heredia, R.R., Cieślicka, A.B. (Eds.), Bilingual Figurative Language Processing. Cambridge University Press, pp. 208–244.

Conklin, K., Schmitt, N., 2008. Formulaic sequences: are they processed more quickly than nonformulaic language by native and nonnative speakers? Appl. Linguist. 29 (1), 72–89. http://dx.doi.org/10.1093/applin/amn022

Ellis, N.C., 2012. Formulaic language and second language acquisition: zipf and the phrasal teddy bear. Annu. Rev. Appl. Linguist. 32, 17–44. http://dx.doi.org/10.1017/S0267190512000025

Ellis, N.C., Simpson-Vlach, R., Maynard, C., 2008. Formulaic language in native and second language speakers: psycholinguistics, corpus linguistics, and TESOL. TESOL Q. 42 (3), 375–396.

Eyckmans, J., Boers, F., Stengers, H., 2007. Identifying chunks: who can see the wood for the trees? Lang. Forum 33 (2), 85–100 https://cris.vub.be/en/publications/identifying-chunks-who-can-see-the-wood-for-the-trees(093b790a-e2bc-4bf9-b2b1-4f0da180966b).html

Fox, J., 2003. Effect displays in R for generalised linear models. J. Stat. Softw. 8 (15), 1–27. http://dx.doi.org/10.2139/ssrn.271037

García, O., Cieślicka, A.B., Heredia, R.R., 2015. Nonliteral language processing and methodological considerations. In: Heredia, R.R., Cieślicka, A.B. (Eds.), Bilingual Figurative Language Processing. Cambridge University Press, pp. 117–168. http://dx.doi.org/10.1017/CBO9781139342100.009

Gemsbacher, M.A., 1984. Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. J. Exp. Psychol. Gen. 113 (2), 256–281. http://dx.doi.org/10.1016/j.jacc.2007.01.076

Güngör, F., Uysal, H.H., 2020. Lexical bundle use and crosslinguistic influence in academic texts. Lingua 242. http://dx.doi.org/10.1016/j.lingua.2020.102859

Hubers, F., van Ginkel, W., Cucchiari, C., Dijkstra, T., & Strik, H. (2016). How native speakers see the light. Paper Presented at FLaRN 2016. Hubers, F., van Ginkel, W., Cucchiari, C., Strik, H., Dijkstra, T., 2018. Normative data on Dutch idiomatic expressions: native speakers. DANS [Dataset]. http://dx.doi.org/10.17026/dans-zjx-hnsk

Hubers, F., Cucchiari, C., Strik, H., Dijkstra, T., 2019. Normative data of Dutch idiomatic expressions: subjective judgments you can bank on. Frontiers in Psychology 10, 1–15. http://dx.doi.org/10.3389/fpsyg.2019.01075

Irujo, S., 1986a. Don’t put your leg in your mouth: transfer in the acquisition of idioms in a second language. TESOL Q. 20 (2), 287–304.

Irujo, S., 1986b. A piece of cake: learning and teaching idioms. ELT J. 40 (3), 236–243. http://dx.doi.org/10.1093/elt/40.3.236

Kecskes, I., 2007. Formulaic language in English Linguia Franca. In: Kecskes, I., Hom, L.R. (Eds.), Explorations in pragmatics: linguistic, cognitive and intercultural aspects. Mouton de Gruyter, pp. 191–219.

Keyser, B., Bly, B., 1995. Intuitions of the transparency of idioms: can one keep a secret by spilling the beans? J. Memory Lang. 34 (1), 89–109. http://dx.doi.org/10.1016/0022-2402(95)90106-3

Kim, C., 2016. L2 learners’ recognition of unfamiliar idioms composed of familiar words. Lang. Aware. 25 (1–2), 89–109. http://dx.doi.org/10.1080/09658416.2015.1122025

Kövecses, Z., Szabó, P., 1996. Idioms: a view from cognitive semantics. Appl. Linguist. 17 (3)326–255.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. ImerTest package: tests in linear mixed effects models. J. Stat. Softw. 82 (13). http://dx.doi.org/10.18637/jss.v082.i13

Lauber, F., 1997. The lexical plight in second language reading: words you don’t know, words you think you know, and words you can’t guess. In: Coady, J., Huckin, T. (Eds.), Second Language Vocabulary Acquisition. Cambridge University Press, pp. 20–34–34. http://dx.doi.org/10.1017/CBO9781139524643.004

Lauber, F., 2000. avoidance of idioms in a second language: the effect of L1–L2 degree of similarity. Stud. Linguist. 54 (2), 186–196. http://dx.doi.org/10.1111/1467-9582.00059

Lemhöfer, K., Broersma, M., 2012. Introducing LexTALE: a quick and valid lexical test for advanced learners of English. Behav. Res. Methods 44 (2), 325–343. http://dx.doi.org/10.3758/s13428-011-0146-0

Libben, Titone, 2008. The multidetermined nature of idiom processing. Memory Cognit. 36 (6), 1103–1121. http://dx.doi.org/10.3758/MC.36.6.1103

Liontas, J., 2002. Context and idiom understanding in second languages. EUROSLA Yearbook 2, 155–185. http://dx.doi.org/10.1075/eurosla.2.11lion

Lo Martire, R., 2017. rel: Reliability Coefficients. (R package version 1.3.1). https://cran.r-project.org/package=rel

Malt, B.C., Eiter, B., 2004. Even with a green card, you can be put out to pasture and still have to work: non-native intuitions of the transparency of common idioms. Memory Cognit. 32 (6), 896–904 http://www.ncbi.nlm.nih.gov/pubmed/15673178

Nordmann, E., Cleland, A.A., Bull, R., 2014. Familiarity breeds dissent: reliability analyses for British-English idioms on measures of familiarity, meaning, literalness, and decomposability. Acta Psychol. 149, 87–95. http://dx.doi.org/10.1016/j.actpsy.2014.03.009

Nordmann, E., Jambazova, A.A., 2017. Normative data for idiomatic expressions. Behav. Res. Methods 49, 198–215. http://dx.doi.org/10.3758/s13428-016-0705-5

Nunberg, G., 1978. The Pragmatics of Reference. Bloomington Indiana University Linguistics Club, Bloomington.

Oostdijk, N., Reynaert, M., Hoste, V., Schuurman, I., 2013. The construction of a 500-million-word reference corpus of contemporary written Dutch. Essential Speech and Language Technology for Dutch 219–247.

Paivio, A., 1986. Mental Representations: A Dual Coding Approach. Oxford University Press.

Pawley, A., Syder, F.H., 1983. Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In: Richards, J.C., Schmidt, R.W. (Eds.), Language and Communication, Longman, pp. 191–225.

Peters, E., 2012. Learning German formulaic sequences: the effect of two attention-drawing techniques. Lang. Learn. J. 40 (1), 65–79. http://dx.doi.org/10.1093/ijl/12085736.2012.685224

Qualtrics, 2005. Qualtrics. http://www.qualtrics.com

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing https://www.r-project.org/
Ramonda, K., 2019. The role of encyclopedic world knowledge in semantic. Engl. Lang. Linguist. 23 (1), 31–53. http://dx.doi.org/10.1017/S136074317000284
Rietveld, T., van Hout, R., 1993. Statistical Techniques for the Study of Language and Language Behaviour. Mouton de Gruyter.
Sadowski, M., 2005. A dual coding view of vocabulary learning. Read. Writ. Q. 21 (3), 221–238. http://dx.doi.org/10.1080/10573560590949359
Sawilowsky, S.S., 2009. New effect size rules of thumb. J. Modern Appl. Stat. Methods 8 (2), 597–599. http://dx.doi.org/10.22237/jmasm/1257035100
Sliyano-Chantruria, A., Spina, S., 2015. Investigation of native speaker and second language learner intuition of collocation frequency. Lang. Learn. 65 (3), 533–562. http://dx.doi.org/10.1111/lang.12125
Skoufaki, V., van Rietveld, T., 2013. The role of encyclopedic world knowledge in semantic. Appl. Linguist. 23 (1), 31–53. http://dx.doi.org/10.1017/S136074317000284
Steinel, M.P., Hulstijn, J.H., Steinel, W., 2007. second language idiom learning in a paired-associate paradigm: effects of direction of learning, direction of testing, idiom imageability, and idiom transparency. Stud. Second Lang. Acquisit. 29, 449–484.
Tinsley, H.E.A., Brown, S.D., 2000. Handbook of Applied Multivariate Statistics and Mathematical Modeling. Academic Press.
Tinsley, H.E.A., Weiss, D.J., 1975. Interrater reliability and agreement of subjective judgments. J. Counsel. Psychol. 22 (4), 358–376. http://dx.doi.org/10.1037/h0076640
Titone, D., Columbus, G., Whitford, V., Mercier, J., Libben, M., 2015. Contrasting bilingual and monolingual idiom processing. In: Heredia, R.R., Cieslicka, A.B. (Eds.), Bilingual Figurative Language Processing. Cambridge University Press. pp. 171–207.
Titone, D., Connine, C.M., 1994. Descriptive norms for 171 idiomatic expressions: familiarity, compositionality, predictability, and literality. Metaphor Symbol. Activity 9 (4), 247–270.
van Ginkel, W., Hubers, F., Cucchiarini, C., Dijkstra, T., & Strik, H. (2016). Norming Studies for Idiom Processing: Native and Non-Native Benchmarks. Paper Presented at FLaRN 2016.
Wray, A., 2000. Formulaic sequences in second language teaching: principle and practice. Appl. Linguist. 21 (4), 463–489. http://dx.doi.org/10.1093/aplin/21.4.463
Wray, A., 2002. Formulaic Language and the Lexicon. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511519772
Yorio, C.A., 1989. Idiomaticity as an indicator of second language proficiency. In: Hytenstam, K., Obler, L.K. (Eds.), Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss. Cambridge University Press. pp. 55–72. http://dx.doi.org/10.1017/CBO97805111780.005
Zyzik, E., 2011. Second language idiom learning: the effects of lexical knowledge and pedagogical sequencing. Lang. Teach. Res. 15 (4), 413–433. http://dx.doi.org/10.1177/13621688114142025