Lexique
ISSN : 2804-7397
Éditeur : Université de Lille

37 | 2025
Le lexique préfabriqué des interactions : des descriptions linguistiques et outillées à l'application didactique

Learning money idioms in L2 English: A window into proficiency and error patterns

Apprendre les expressions idiomatiques liées à l'argent en anglais langue seconde : une fenêtre sur les compétences et les types d'erreurs

Elena Gavruseva

https://www.peren-revues.fr/lexique/2080

DOI : 10.54563/lexique.2080

Référence électronique
Elena Gavruseva, « Learning money idioms in L2 English: A window into proficiency and error patterns », *Lexique* [En ligne], 37 | 2025, mis en ligne le 01 décembre 2025, consulté le 27 février 2026. URL : https://www.peren-revues.fr/lexique/2080

Droits d'auteur
CC-BY

Learning money idioms in L2 English: A window into proficiency and error patterns

Apprendre les expressions idiomatiques liées à l'argent en anglais langue seconde : une fenêtre sur les compétences et les types d'erreurs

Elena Gavruseva

PLAN

1. Introduction
2. Background
3. The Study
   3.1. Participants
   3.2. Materials
   3.3. Results
      3.3.1. Accuracy rates by self-reported proficiency level
      3.3.2. Idiom-based learner groups
      3.3.3. Idiom difficulty ranking
      3.3.4. External factors
4. Discussion
5. Conclusion

NOTES DE LA RÉDACTION

Reçu : Mars 2025 / Accepté : Septembre 2025
Mise en ligne : Décembre 2025

NOTES DE L'AUTEUR

I am grateful to the participants in the experiment, the two anonymous reviewers for their extensive and constructive criticism, and the editors of this special issue, Carole Etienne and Alexis Ladreyt, for their professional guidance and feedback. I also thank Knute Carter and Peter Kortenkamp for their assistance with statistical analyses. I thank Julia Andrews, Anna Dyer, Jacee Cho, Elena Kallestinova, Silvia Lattova, Olga Meerson, and Timur Yarovinsky for their assistance in recruiting participants for this study. This research was supported by the Investment in Strategic Priorities Grant from the Provost's Office at the University of Iowa, awarded to the author, and was approved by the University of Iowa Human Subjects Office. All remaining errors are my own.

TEXTE

1. Introduction

1 Idiomatic expressions represent a distinctive and often elusive component of second language (L2) learning. Unlike collocations used with literal meanings, idioms are typically figurative, culturally embedded, and highly variable in usage, which makes them a persistent challenge for L2 learners. As a result, idiomatic competence is often viewed as a marker of advanced proficiency levels and a key indicator of advanced or native-like fluency (Abel, 2003; Bell, 2017; Cenoz, Hufeisen & Jessner, 2003; Guo, 2014; Irujo, 1986, 1993; Kellerman, 1986; Laufer, 2000; Lin, 2016; Van Ginkel & Dijkstra, 2020; Wu, Chen & Huang, 2006). Within this broader domain, idioms belonging to specialized lexical fields – such as the pecuniary lexicon – offer a unique and underexplored opportunity for assessing language development. Money-related idioms (e.g. *bring home the bacon*, *turn on a dime*, etc.) are densely packed with cultural connotations and context-dependent meanings, which make them especially sensitive to variation in language experience, quality of input, input frequency, and interpretive strategies among non-native speakers. Given the prevalence of money idioms in the English lexicon, this semantic field offers a promising resource for investigating their potential as an instrument of language assessment.

2 Despite their complexity, certain money idioms may be acquired early by L2 learners due to factors such as semantic transparency, crosslinguistic overlap, or high frequency in (social) media and everyday conversation. Others, however, may remain opaque and resistant to interpretation, often due to culture-specific references or metaphorical density (e.g. *It is burning a hole in my pocket*). Understanding such expressions requires cognitive flexibility, suppression of literal meanings, and access to culturally shaped metaphorical reasoning––skills that extend beyond rote vocabulary memorization and syntactic competence (Charteris-Black, 2002; Nacey, 2013). This article explores the hypothesis that this variation makes idioms a valuable diagnostic tool for distinguishing not only between non-

native and native speakers, but also between proficiency levels among L2 learners.

3 In this study, we define idiomatic competence as speakers' ability to recognize, interpret, and appropriately use idiomatic expressions––multi-word units whose meaning is not fully derivable from the literal meanings of their components. This competence encompasses knowledge of both form and meaning, including expressive, cultural, and pragmatic functions of idioms; yet, despite its richness, idiomatic competence has rarely been used systematically as a metric for assessing proficiency. We address this gap by examining how L1-Slavic/L2-English learners acquire money-related idioms through a recognition/comprehension task that targets the receptive dimension of their idiom knowledge.

4 Our main goal is to investigate whether knowledge of money-related idioms, as measured by test scores, can serve as an indicator not only of overall self-rated English proficiency but also of specific language abilities––namely, speaking, understanding, and reading. More broadly, we explore whether idiom performance can serve as a marker of non-native status. A secondary goal is to examine how accuracy on individual idioms varies across proficiency levels in order to identify which idioms tend to be acquired early and which are acquired later. This work also considers how idiom recognition is modulated by learners' proficiency level, as well as by external, experiential factors such as education level, age of onset of English learning, age of arrival (immigration), and length of residence in the US.

5 This article is organized as follows. Section 2 discusses research on idioms and their relationship with language processing and language proficiency. Section 3 introduces the experiment and reports the findings. Section 4 discusses the results. Section 5 concludes the article.

2. Background

6 Idioms can be thought of as long words––fixed expressions that function as single lexical units––even though they vary in internal compositionality (sentences vs. phrases). Some idioms retain partial

semantic transparency, while others do not (Bally, 1921; Mel'čuk, 2023). For example, the popular expression among American youth She ate, no crumbs (meaning 'She did a perfect job') allows for an imagery-based mapping between its literal and figurative meanings: someone who thoroughly clears her plate is metaphorically seen as having performed exceptionally well. Other idioms, such as go cold turkey––meaning to suddenly and completely stop engaging in an addictive behavior––exhibit greater semantic opacity. This variability in transparency, along with idiom familiarity and literal plausibility, has been a central focus in psycholinguistic research, particularly in studies of real-time idiom processing in monolinguals (e.g. Gibbs, 1985; Katz, 2024; among many others)[1].

7 Studies on bilingual idiom processing typically compare native speakers with L2 learners or examine how bilinguals process idioms in their L1 versus L2. Research shows that native speakers process idioms and formulaic expressions faster than non-native speakers (Carrol & Conklin, 2017; Siyanova-Chanturia, Conklin & Schmitt, 2011; among others). The picture becomes more nuanced in studies on idiom processing within bilinguals (e.g. Cieślicka, 2006, 2013, 2015; Isobe, 2011; Titone, Columbus, Whitford, Mercier & Libben, 2015; among others). For instance, Cieślicka (2013) found that idiom (non)decomposability plays a key role in determining L1–L2 processing similarities/differences in L1-Polish/L2-English bilinguals. Studies that consider cross-linguistic influence suggest that idioms shared across languages––so-called congruent expressions––can offer a processing advantage in L2, likely due to their prior lexicalization in L1 (Carrol, Conklin & Gyllstad, 2016; Titone, Columbus, Whitford, Mercier & Libben, 2015; Wolter & Gyllstad, 2013)[2].

8 Interpretation of processing results is often limited by how L2 proficiency is treated. Though commonly measured, proficiency is rarely analyzed as an independent variable, with non-native participants often grouped together regardless of being foreign language or L2 learners. These groups, typically defined by self-reports or standardized tests (e.g. TOEFL), often include varied proficiency levels and diverse L1 backgrounds. Treating them as homogeneous may obscure some of the factors modulating processing patterns, like L1 influence and proficiency-based strategies.

9 Relatively few studies have explored how idiomatic competence interacts with, or reflects, overall L2 proficiency levels––or specific subcomponents such as speaking, reading, etc. While traditional L2 assessment methods tend to focus on grammar and general vocabulary, idiomatic knowledge (particularly of culturally embedded, metaphor-based expressions) remains underexplored as a diagnostic tool or marker of language proficiency. Vanderniet (2015) investigated the relationship between idiomatic knowledge and proficiency in a study of 72 second/foreign-language learners, primarily from Spanish, Korean, and Chinese L1s. Participants were grouped into six proficiency levels based on scores from their university's standardized Language Acquisition Test (LAT). Each participant also had LAT component scores for reading, writing, and speaking. They completed an idiom comprehension task with 12 items (e.g. throw a fit) in multiple-choice format. Results showed strong positive correlations between idiom comprehension and both overall LAT and speaking scores, but no significant correlation with reading or writing. Vanderniet also administered a 24-item idiom test to 340 English learners via Amazon Mechanical Turk, who self-rated their proficiency on a 1–10 scale. Pearson's test again showed a significant positive correlation between idiom test scores and self-assessed proficiency. Proficiency effects are further supported by studies of L1-Arabic/L2-English learners (e.g. Aljabri, 2013).

10 We expand on previous research by investigating how idiom recognition varies across proficiency levels and how L2 learners from Slavic L1 backgrounds compare to native speakers in their knowledge of money-related idioms. Specifically, we explore whether idiom knowledge within a single semantic domain can serve as a reliable and practical indicator of overall L2 proficiency. To address this, we examine accuracy patterns and the relative difficulty of individual idioms across proficiency levels. In addition, we consider the influence of external factors such as age of onset (first contact with L2), age of arrival, educational background, and length of residence on idiom accuracy. The following research questions guide our investigation:

1. How do accuracy rates on money idioms vary across L2 learners with different self-reported proficiency levels?

2. Which money idioms present the greatest and least difficulty for L2 learners, and how does their relative difficulty ranking shift across proficiency levels?

3. To what extent do accuracy rates on money idioms remain consistent across different proficiency levels, and what patterns of overlap emerge?

11 By empirically testing the hypothesis that idiom knowledge reflects broader linguistic competence, this study contributes to both assessment theory and applied linguistics. If validated, the proposed idiom-based framework could offer a more fine-grained approach to measuring L2 lexical proficiency––one that reflects not only linguistic knowledge but also deeper semantic knowledge and metaphorical reasoning.

3. The Study

12 This section introduces an experimental study designed to investigate the relationship between idiomatic knowledge and self-reported proficiency in L2. We begin by describing the participants and their language backgrounds, followed by an overview of the materials and procedures used in the study. Finally, we present the results, highlighting accuracy patterns and proficiency-based differences in error trends.

3.1. Participants

13 The study included 45 native American English speakers (ages 21–74, M = 35, SD = 12; 30 females, 12 males, 3 non-binary) and 52 L2-English learners. Among the native speakers, 14 held a BA/BS degree, 12 an MA/MS, 5 a PhD; 2 were pursuing a BA/BS, 4 an MA/MS; 4 had completed high school, and 4 selected 'other'. Native speakers were recruited via Reddit platform where they accessed an anonymous survey link (https://www.reddit.com/r/SampleSize). A brief biographical questionnaire was embedded within the idiom survey. The participants completed both parts of the survey on their personal devices.

14 All L2 participants were native speakers of a Slavic language (primarily Ukrainian and Russian) and had learned English subsequently to their L1. They were recruited through the author's personal and professional networks. All had some experience living and learning English in the United States and therefore were all second language (ESL) learners. Most (45/52) began learning English in their country of birth, typically through formal classroom instruction. Consequently, their language learning profiles were predominantly mixed, combining instructed foreign language learning with subsequent immersion in an English-speaking environment. Testing was conducted entirely online, with participants completing the survey on their personal devices. Each participant received two anonymous links: one for the biographical survey and another for the idiom test. The biographical survey was a shortened version of the Language Experience and Proficiency Questionnaire––LEAP-Q (Marian, Blumenfeld, & Kaushanskaya, 2007)[3].

15 The survey collected information on participants' native language(s), demographic details, and educational background. Participants also self-rated their English proficiency using the following categories: near-native, advanced, high intermediate, intermediate, low intermediate, or beginner, and assessed their speaking, understanding, and reading skills in English separately on a 10-point Likert scale. Most of the learners (49/52) were residing in the U.S. at the time of the study. Table 1 summarizes their demographic and language experience profiles.

	Age at time of study	**Age of onset**	**Age of arrival**	**Years in U.S.**	**Speaking**	**Understanding**	**Reading**
Native-like (n = 21)							
Range	26–66	4–25	15–36.4	1–34.6	8–10	6–10	8–10
Mean	47.5	12	25.7	20.7	9.19	9.52	9.52
SD	12.7	6.1	5.6	11.6	0.98	0.93	0.68
Advanced (n=21)							

Range	22–74	5;5–30	17–43	2–46.5	4–10	3–10	5–10
Mean	50.1	11.5	26.2	23.9	8.19	8.71	9.14
SD	16.3	5.6	6.6	13.3	1.44	1.62	1.20
Intermediate (n = 10)							
Range	39–69	8–35	27–37	25–39	5–7	4–8	4–9
Mean	58.5	18.2	31.1	25.4	6.2	6.7	6.8
SD	8.5	9.9	3.8	13.1	1.03	1.49	1.62

Table 1. Language experience and background information for L2 participants

16 *Native-like Group* included 21 participants (18 females, 3 males), with 20 L1-Russian speakers and one bilingual (Russian/Ukrainian). All acquired English after their L1 (age of onset: 4–25, M = 12); 16 began learning English in their birth country, 5 in the US. Seventeen reported Russian as dominant, 3 reported English, and 1 listed Russian, English, and French equally. Education: 10 Master's, 7 PhDs (2 in progress), 3 BAs, and 1 'other'.

17 *Advanced Group* included 21 participants (15 females, 6 males); 17 were L1-Russian speakers and 4 were Russian/Ukrainian bilinguals. All acquired English after their L1 (age of onset: 5;5–30, M = 11.5); 20 began learning English in their birth country, 1 in the US. Sixteen reported Russian as dominant, 1 reported Ukrainian/Russian, and 4 reported Russian/English as equally dominant. Education: 9 Master's (2 in progress), 7 PhDs (2 in progress), 3 BAs, and 2 with professional training.

18 *Intermediate Group* included 10 participants (8 females, 2 males), all L1-Russian. Nine began learning English in their country of birth, one in the U.S. All acquired English after L1, with onset ages from 8 to 35 years (M = 18.2). Nine were residing in the U.S. at the time of the study. All listed Russian as their dominant language. Education: 4 Master's, 2 PhDs, 4 Bachelor's.

19 Table 2 presents the results of Wilcoxon rank-sum tests comparing self-rated language abilities across the three L2 groups. P-values were adjusted using the Benjamini-Hochberg procedure to control for multiple comparisons. These results offer insight into how L2

learners with different self-assigned proficiency levels perceive their specific language skills.

Ability Rating	Comparisons		W value	p-adj
Speaking	Native-like vs.	Adv	298	p < .01
Speaking	Native-like vs.	Int	174	p < .0001
Speaking	Adv vs.	Int	153	p < .003
Understanding	Native-like vs.	Adv	282	p < .02
Understanding	Native-like vs.	Int	174	p < .0001
Understanding	Adv vs.	Int	155	p < .003
Reading	Native-like vs.	Adv	236	p = 0.289
Reading	Native-like vs.	Int	172	p < .0002
Reading	Adv vs.	Int	162	p < .001

Table 2. Wilcoxon pairwise comparisons by self-rated language ability levels

20 For speaking ability, significant differences were found between all three groups. Native-like learners rated their speaking ability significantly higher than both Advanced (Adv) ($p < .01$) and Intermediate (Int) ($p < .0001$), and Advanced learners rated themselves significantly higher than Intermediate ($p < .003$). A similar pattern held for understanding ability.

21 Reading ability also showed between-group differences, with Intermediate learners rating their reading significantly lower than both Advanced ($p < .001$) and Native-like ($p < .0002$). However, no significant difference was found between the Native-like and Advanced learners. This pattern may suggest that while reading helps distinguish lower proficiency learners from higher ones, it may be less effective in differentiating among more advanced, adjacent levels of proficiency, at least based on self-assessments.

22 Overall, these analyses demonstrate that learners' self-rated oral abilities (speaking and understanding) align more consistently with their broader self-assigned proficiency category than reading ability. This finding will inform our interpretation of idiom accuracy patterns in the Results section.

3.2. Materials

23 The study was conducted using an online test designed in Qualtrics. The test included 36 items: 12 money idioms as experimental targets and 24 distractor items. The idioms and their meanings are given in Table 3.

Idiom	**Meaning**
1. burning a hole in my pocket	feeling an urge to spend money quickly
2. if I had a nickel for every time I've heard	used to emphasize frequent occurrences
3. IOUs (I owe you)	a written acknowledgement of debt
4. bring home the bacon	earn money to support oneself or family
5. not buying it	not believing or accepting something as true
6. in for a penny, in for a pound	once committed to something, commit fully
7. won't break the bank	affordable, not too expensive
8. haven't spent one red cent	haven't spent any money at all
9. turn on a dime	change very quickly
10. bet your bottom dollar	be absolutely sure about something
11. cost a pretty penny	very expensive
12. sugar daddy	an older, wealthy man who gives money or gifts to a younger person, often in exchange for companionship or romantic relationship

Table 3. Money idioms and their meanings

24 *Idiom selection procedure*: To compile the list of money idioms for the test, two native speakers of American English independently generated idioms they commonly used or heard. Their lists were combined, with overlapping items counted as the same, resulting in a total of 37 idioms. Twelve idioms were then randomly selected for inclusion in the test [4]. Among the idioms selected, sugar daddy has a partial equivalent in Russian (papik––a suffixed diminutive of papa, meaning *daddy*), though it lacks the explicit modifier *sugar*. Another idiom––it will cost you a pretty penny––has a loose equivalent in Russian that conveys a similar meaning (*eto vl'etit t'eb'e v kopejechku*; lit. 'this will

fly into your penny'), but differs significantly in form (it uses a different verb, no adjective equivalent to pretty, and refers to kopejechku––a different coin denomination). Bring home the bacon also has a partial equivalent in Russian: zarabatyvat' na khleb (lit. 'to earn money for bread'). While both idioms refer to providing for one's household, they differ in lexical content––bread instead of bacon––and in structure, with the Russian version using the verb zarabatyvat' rather than a construction *bring home*. The idiom that shows the greatest degree of lexical/conceptual overlap is haven't spent one red cent, which closely corresponds to the Russian expression ni kopejki ne potratil (lit. 'not a cent not spent'). While there is a difference in object-verb word order and the absence of a modifying adjective, the idioms are conceptually equivalent. The remaining idioms showed no overlap with the learners' L1s in lexical items, syntactic structure, or conceptual meaning.

25 *Stimuli design*: To construct experimental sentences, we searched the Corpus of Contemporary American English (COCA) for naturalistic examples of these idioms, randomly selecting four sentences per idiom. This process yielded 48 sentences (12 idioms × 4 sentences each). From this set, 12 sentences were randomly selected as experimental items for the test[5]. The idiom test included 12 multiple-choice questions, divided into three formats, all testing idiom recognition/comprehension:

- *Fill-in-the-blank*: Six items required participants to select the correct word from five multiple-choice options to complete a sentence. An example is given in (1).

- *Meaning differentiation*: Four items presented three sentences/phrases, and participants were asked to identify the one that differed in meaning from the other two. An example is given in (2).- *Word meaning selection*: Two items focused on a specific word, asking participants to choose a synonym from multiple-choice options.

(1) *Fill-in-the-blank item:*
The problem is one we can solve, and it won't __ the bank.

(a) empty (b) clear (c) hurt (d) break (e) hammer

(2) *Meaning differentiation item:*
Which sentence does not share meaning with the other two?

(a) I am not into it.
(b) I don't buy it.
(c) I don't believe it.

26 Distractor items had similar multiple-choice formats [6]. An example is given in (3):

(3) The word *callow* is similar to which of the words below?
(a) green (b) shallow (c) pillow (d) callous (e) canny

27 All test items (experimental and distractors) were randomized for each participant. The response options within the multiple-choice menu were also randomized. Before completing the test, participants were presented with a consent form and detailed instructions emphasizing that they should rely on their first instinct, consider all options carefully, and not consult dictionaries or any online resources [7].

3.3. Results

28 We first present the accuracy rates of idiom test items based on self-reported proficiency levels. One of the 12 test items (*sugar daddy*) was excluded due to an inadvertent error in the survey [8]. The statistical analyses were performed using R (R Core Team, 2025) and are based on the participants' responses to 11 idiom test items.

3.3.1. Accuracy rates by self-reported proficiency level

29 Figure 1 presents a box plot illustrating the distribution of correct responses by self-reported proficiency level.

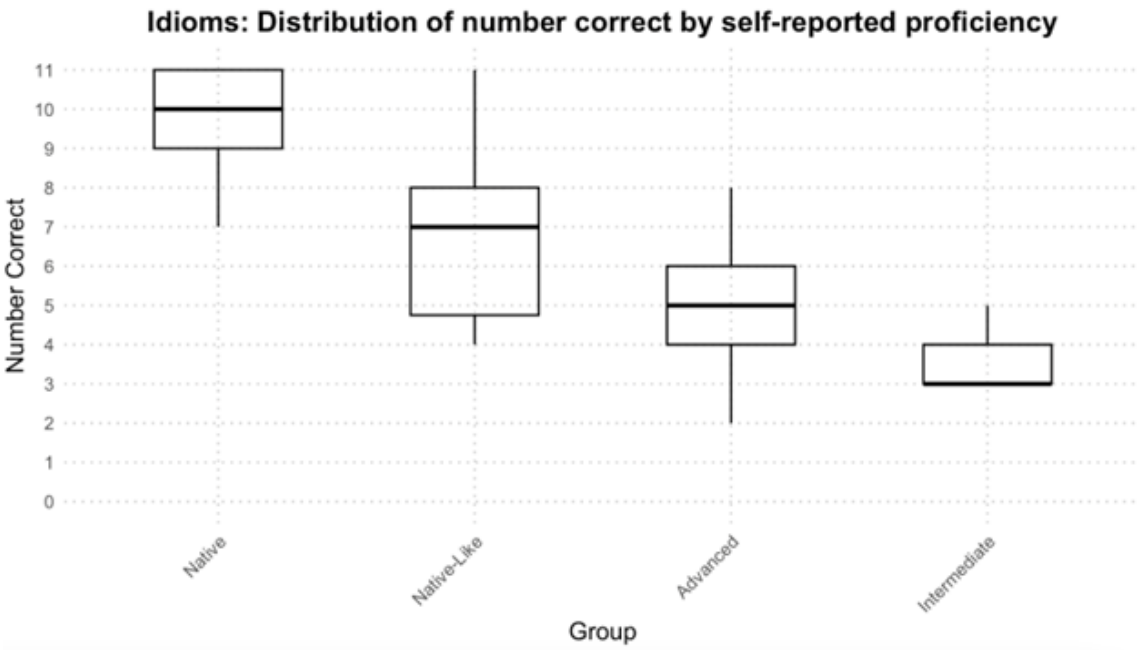

Figure 1: Distribution of number correct by self-reported proficiency

30 The box plot illustrates the distribution of correct idiom responses across four groups: Native, Native-Like, Advanced, and Intermediate. The native speakers exhibit the highest median score of 10, with some variability (7–11 range), indicating overall consistency in their knowledge of idioms. The Native-Like group, while performing lower than the native speakers, with the median score of 7, shows greater variability range (4–11). The Advanced group has a further reduced median score of 5, with a wide variability range (2–8), also indicating large individual differences in idiom knowledge. The Intermediate group has the lowest median score of 3 and a small interquartile range, suggesting that most participants in this group struggle with idioms. Overall, the data demonstrate a clear proficiency-related trend, with higher self-reported proficiency aligning with greater idiom knowledge. Additionally, greater variability in the Native-Like and Advanced groups suggests that idiomatic knowledge does not develop uniformly, with some learners performing closer to native speaker levels while others lag behind.

31 Having established the general patterns of idiom accuracy across groups, we now turn to statistical analysis to evaluate these differences. Shapiro-Wilk and Levene's tests were first conducted to assess normality and homogeneity of variance. As three out of four groups

violated the normality assumption and homogeneity of variance was not met, a standard ANOVA was deemed inappropriate; therefore, we conducted Welch's ANOVA to examine overall group differences. The analysis showed a statistically significant effect ($F = 91.659$, $df = 3$, $p = 7.022e-14$, $p < 0.001$), indicating a meaningful difference in idiom accuracy across proficiency levels.

32 Next, we performed post-hoc comparisons using the Games-Howell test to determine which groups differed from each other. The results are presented in Table 4.

Comparison	**Mean difference**	**95% CI (Lower, Upper)**	**p-value**	**Significance**
Native (n=45) vs. Native-Like (n=21)	-0.280	(-0.417, -0.142)	5.98e-5	p < .0001
Native vs. Advanced (n=21)	-0.416	(-0.508, -0.324)	6.39e-12	p < .0001
Native vs. Intermediate (n=10)	-0.596	(-0.741, -0.451)	1.25e-6	p < .0001
Native-Like vs. Advanced	-0.136	(-0.290, 0.0169)	0.095	ns (not significant)
Native-Like vs. Intermediate	-0.316	(-0.497, -0.135)	3.65e-4	p < .001
Advanced vs. Intermediate	-0.180	(-0.336, -0.0240)	0.021	p < 0.02

Table 4. Pairwise comparisons (Games-Howell post-hoc test)

33 As shown in Table 4, there was a significant difference in idiom accuracy rates ($p < .0001$) between native speakers (89%) and all L2 groups. Turning to the pairwise comparisons within the L2 sample, there was no significant difference between the Native-Like and the Advanced learners (61% vs. 47%, $p < 0.095$). However, there were statistically significant differences between the Native-Like and the Intermediate speakers (61% vs. 29%, $p < 0.001$), as well as between the Advanced and the Intermediate speakers (47% vs. 29%, $p < .02$). These findings suggest that the most notable differences occur between native speakers and all L2 groups. Within the L2 sample, idiom accuracy helps distinguish lower-proficiency learners from higher ones, but it appears to be less effective in differentiating among more advanced proficiency levels––at least as defined by broad self-rated categories.

34 To further assess the relationship between learners' self-perceived language abilities and their idiom test performance, Spearman rank correlations were computed for self-ratings in speaking, understanding, and reading. The results are presented in Table 5.

Rating	Spearman ρ	p-value
Speaking	0.481	0.0001
Understanding	0.282	0.05
Reading	0.266	0.06

Table 5. Spearman rank correlations between self-rated language abilities and idiom test scores

35 The results revealed a statistically significant positive correlation between self-rated speaking ability and idiom test scores ($p < .0001$), indicating that L2 learners who rated themselves more highly in speaking tended to perform better on the idiom test. A weaker but still significant correlation was observed for self-rated understanding ($p < .05$). In contrast, self-rated reading showed a non-significant trend, indicating that confidence in reading ability was not associated with idiom test performance in this L2 sample.

3.3.2. Idiom-based learner groups

36 This section focuses on differences and similarities in idiom proficiency across learner groups. In Section 3.3.1, we noted that the Native-Like and Advanced groups exhibited substantial individual variation and broad idiom score ranges, which may help explain the lack of statistically significant differences between them in the Games-Howell analysis. This variability also suggests that broad self-reported proficiency labels may not reliably reflect differences in idiom knowledge in advanced L2 speakers. Consequently, idiom test performance itself may offer a more informative basis for regrouping participants, yielding internally consistent groups whose score variability (measured by standard deviation) is more comparable to that of the native speaker group in this study. This reclassification will reduce within-group dispersion in test scores, allowing for clearer contrasts across groups and more precise analysis of developmental patterns and learner errors.

37 Using the data from Figure 1, we will now redefine L2 proficiency groups based on their idiom test performance. Specifically, we will use the median scores and the top 25% of participants (represented by the upper whiskers) to create three new groups. The first group is based on the median score of 7, including all L2 participants who scored 7 or higher – this threshold also aligns with the lowest score observed among native speakers. The second group is defined by the median score of 5, comprising L2 participants who scored at 5 and up, but below 7. The third group consists of all remaining L2 participants, who scored below 5. This approach enables us to reclassify non-native speakers into new, idiom-based proficiency levels, designated as Group 1 (highest), Group 2 (second highest), and Group 3 (lowest).

38 Table 6 presents mean accuracy rates as the percentage of correct responses for Group 1 (G1), Group 2 (G2), and Group 3 (G3), in comparison to the Native group.

Group	Accuracy rate	SD	Min	Max
Native speakers (n=45)	.89	.09	.64	1
G1 (n=16)	.74	.13	.64	1
G2 (n=17)	.49	.05	.46	.55
G3 (n=19)	.31	.09	0	.36

Table 6. Mean accuracy rates and standard deviations for idiom-based groups

39 Welch's ANOVA was used to examine overall group differences. The analysis showed a statistically significant effect, indicating a meaningful difference in idiom accuracy across the newly formed groups ($p < 0.001$). Next, we performed post-hoc comparisons, using the Games-Howell test to determine which groups differed from each other. The results are presented in Table 7.

Comparison	Mean difference	95% CI (Lower, Upper)	Significance
Native vs. G1	+ 0.149	(0.045, 0.254)	$p < .01$
Native vs. G2	+ 0.398	(0.349, 0.447)	$p < .0001$
Native vs. G3	+ 0.583	(0.514, 0.651)	$p < .0001$
G1 vs. G2	- 0.248	(-0.351, -0.146)	$p < .0001$
G1 vs. G3	- 0.433	(-0.544, -0.322)	$p < .0001$
G2 vs. G3	- 0.185	(-0.251, -0.119)	$p < .0001$

Table 7. Pairwise comparisons (Games-Howell post-hoc test)

40 The pairwise comparisons reveal significant differences in idiom accuracy between L2 groups, as well as between L2 and native speakers. The contrast between the Native group and all L2 groups is quite pronounced, with the differences observed between Native and G3 speakers ($p < .0001$), as well as between Native and G2 speakers ($p < .0001$). While the difference between Native and G1 speakers is smaller ($p < .01$), it remains statistically significant, suggesting that even highly proficient L2 speakers do not completely match Native-level idiom accuracy. Among L2 groups, G1 significantly outperforms both G2 ($p < .0001$) and G3 ($p < .0001$). Finally, G2 also maintains a meaningful advantage over G3 ($p < .0001$), confirming that idiom recognition develops progressively with proficiency.

41 To examine whether self-perceived language abilities differ among the L2 learners, we compared self-rated speaking, understanding, and reading abilities across the three idiom-based groups. Table 8 reports the results of Wilcoxon rank-sum tests (Benjamini-Hochberg adjusted).

Rating	**Comparisons**		**W**	**p-adj**
Speaking	G1 vs.	G2	159	0.07
Speaking	G1 vs.	G3	221	0.02
Speaking	G2 vs.	G3	166	0.400
Understanding	G1 vs.	G2	147	0.202
Understanding	G1 vs.	G3	189	0.202
Understanding	G2 vs.	G3	152	0.760
Reading	G1 vs.	G2	140	0.343
Reading	G1 vs.	G3	180	0.343
Reading	G2 vs.	G3	156	0.638

Table 8. Wilcoxon pairwise comparisons by self-rated language ability (G1, G2, G3)

42 Speaking ability was the only domain to show statistically significant differences among the idiom-based groups. G1 rated their speaking ability significantly higher than G3 ($p < 0.02$); the difference with G2 was only marginally significant ($p < 0.07$). No significant difference was found between the adjacent groups G2 and G3. This suggests that high idiom accuracy aligns more closely with stronger self-perceived speaking skills.

43 In contrast, understanding and reading abilities did not significantly differ across any of the idiom-based groups. This lack of difference may suggest that learners' receptive skills (particularly in reading) are less strongly associated with idiom test performance than their productive oral abilities.

3.3.3. Idiom difficulty ranking

44 In this section, we present an analysis of idiom difficulty based on accuracy rates across participant groups. We first present tables ranking idioms from most to least difficult, offering a clear view of how different expressions were understood by participants. Particular attention is given to idioms that proved challenging for the L2 groups. Table 9 summarizes accuracy rates and standard deviations–shown in parentheses–in the Native group.

turn on a dime	pretty penny	the bacon	break the bank	bottom dollar	not buy it	hole in my pocket	if I had a nickel	in for a penny	IOUs	one red ce
1.000 (0)	1.000 (0)	0.98 (0.15)	0.96 (0.21)	0.96 (0.21)	0.93 (0.25)	0.89 (0.32)	0.87 (0.34)	0.87 (0.34)	0.73 (0.45)	0.60 (0.49)

Table 9. Individual idiom accuracy rates in the Native group

45 The accuracy rates reveal a clear hierarchy of idiom familiarity. Idioms such as turn on a dime and pretty penny were recognized with perfect accuracy across all participants. Slightly lower but still near-ceiling accuracy was observed for *bring home* the bacon, *won't* break the bank, and *bet your* bottom dollar, suggesting strong familiarity with these expressions.

46 As accuracy rates decline, greater individual variation emerges. Idioms like not buy it, *burning a* hole in my pocket, and if I had a nickel maintain high accuracy but with slightly larger standard deviations, indicating that while most native speakers recognize them, a small subset may be less familiar. Idioms such as IOUs and one red cent show considerably lower accuracy and the largest variability, suggesting they may be less familiar or even outdated for some speakers. Table 10 summarizes individual idiom accuracy rates for G1 group.

break the bank	turn on a dime	not buy it	pretty penny	if I had a nickel	IOUs	in for a penny	hole in my pocket	the bacon	bottom dollar	one red ce
1.00 (0)	1.00 (0)	0.93 (0.26)	0.93 (0.26)	0.87 (0.52)	0.87 (0.52)	0.67 (0.49)	0.67 (0.49)	0.67 (0.49)	0.53 (0.52)	0.40 (0.51)

Table 10. Individual idiom accuracy rates in G1

47 The accuracy rates reveal a stratified pattern of idiom knowledge. The idioms *won't* break the bank and *turn* on a dime achieved perfect accuracy, indicating strong recognition among participants. Not buy it and pretty penny followed closely with 93% accuracy, suggesting that these idioms are also well-known within this group. Accuracy rates begin to decline for if I had a nickel and IOUs (both 87%), followed by in for a penny (in for a pound), *burning a* hole in my pocket, and *bring home* the bacon (all three at 67%). The idioms *bet your* bottom dollar (53%) and *haven't spent* one red cent (40%) ranked the lowest, with a notable increase in variability, suggesting that they are less known in this group.

48 When comparing the top-ranked idioms in G1 with those in the Native group, we observe both similarities and differences. In both groups, turn on a dime achieved perfect accuracy, while pretty penny, *not buy it*, and *won't* break the bank also scored at or near ceiling levels. These four idioms form a subset of the high-accuracy idioms shared between the groups. However, in the Native group, six idioms exceeded the 90% accuracy threshold, including bring home the bacon and bet your bottom dollar. In contrast, these two idioms ranked significantly lower in G1, with accuracy rates of 67% and 53%, respectively, possibly suggesting a gap in idiom exposure between Native and G1 speakers. This comparison highlights the fine-grained differences in idiom knowledge even among the most proficient L2 speakers.

49 Statistical comparisons between the Native and G1 groups on these two idioms (bring home the bacon and bet your bottom dollar) were conducted using a chi-square test, revealing significant differences. Bring home the bacon yielded $p < .05$, while bet your bottom dollar showed an even stronger effect, with $p < .01$. These findings suggest that while G1 participants exhibit near-native proficiency for some idioms, certain expressions may still pose challenges. Table 11 summarizes individual idiom accuracy rates for G2.

not buy it	break the bank	turn on a dime	pretty penny	IOUs	the bacon	in for a penny	one red cent	if I had a nickel	hole in my pocket	bottom doll
0.87 (0.35)	0.87 (0.35)	0.80 (0.41)	0.68 (0.49)	0.60 (0.51)	0.53 (0.52)	0.40 (0.49)	0.27 (0.46)	0.27 (0.46)	0.13 (0.35)	0 0

Table 11. Individual idiom accuracy rates in G2

50 Idiom accuracy rates reveal a pattern that aligns closely with both the G1 and Native groups. The idioms not buy it, won't break the bank, and turn on a dime emerge as best recognized, though none reach the 90% threshold within G2. Instead, they cluster in the 80–89% range, indicating high familiarity but still some room for learning. These three idioms are closely followed by pretty penny (68%). Interestingly, this exact subset of four idioms also forms the highest-performing cluster in the G1 group, where all four exceed 90% accuracy. This suggests a clear transitional cluster that develops in competence from G2 to G1, potentially marking these idioms as early indicators of advancing idiom proficiency. Outside of this cluster, idiom performance becomes more variable, but bottom dollar is particularly striking, with 0% accuracy in G2 (in line with its literal meaning). While its accuracy improves to 53% in the G1 group, it still remains among the least recognized idioms, ranking second to last. Table 12 summarizes idiom accuracy rates for G3.

pretty penny	not buy it	IOUs	break the bank	turn on a dime	the bacon	hole in my pocket	in for a penny	bottom dollar	if I had a nickel	one red ce
0.63 (0.49)	0.58 (0.51)	0.53 (0.51)	0.53 (0.51)	0.42 (0.51)	0.16 (0.38)	0.16 (0.38)	0.16 (0.38)	0.11 (0.32)	0.05 (0.23)	0.05 (0.23)

Table 12. Individual idiom accuracy rates in G3

51 In G3, accuracy rates are noticeably lower, not exceeding 63%. However, an intriguing pattern emerges: the same four idioms – pretty penny, not buy it, won't break the bank, and turn on a dime – are again clustered among the best recognized. An interesting quirk is that G3 performs relatively well on IOUs, with the same accuracy rate as won't break the bank (53%). In fact, IOUs is also among the top five for G2 (in contrast to the Native group where it ranks second to last). Thus, the same cluster of five idioms characterizes both G3 and

G2, suggesting that this idiom cluster may serve as an emergent indicator of growing idiom knowledge.

52 To analyze the consistency of idiom difficulty rankings across proficiency levels, we applied Kendall's Tau, a non-parametric correlation test that measures the strength of ordinal associations between groups. This statistical approach allowed us to determine the degree of agreement in idiom rankings between the native speakers and L2 proficiency groups.

	Native	**G1**	**G2**	**G3**
Native	1.00	0.49	0.25	0.36
G1	0.49	1.00	0.75	0.77
G2	0.25	0.75	1.0	0.61
G3	0.36	0.77	0.61	1.00

Table 13. Kendall's Tau correlations for idiom difficulty rankings across groups

53 The results reveal several key patterns. First, the L2 groups (G1, G2, and G3) show strong agreement with each other in their idiom difficulty rankings, with Kendall's Tau values ranging from 0.61 to 0.77. This suggests that as L2 learners progress in proficiency, the relative difficulty of idioms remains consistent across groups. However, when compared to native speakers, G1 exhibits the highest similarity (τ = 0.49), indicating that advanced learners' rankings align more closely with native competence. In contrast, G2 and G3 diverge more significantly from native speaker rankings, with G2 showing the weakest correlation (τ = 0.25). This suggests that at lower proficiency levels, idiom difficulty rankings deviate more from native speaker patterns, possibly due to differences in exposure or reliance on literal interpretations.

3.3.4. External factors

54 We now turn to the influence of external factors––age of onset, age of arrival, years of residence, and education––to assess their role in idiom proficiency. A series of binomial logistic regression models were conducted to determine how each of these variables, both independently and in combination, impacted the likelihood of correctly answering idiom test items. The results are presented in Table 14.

	Intercept	Coefficients	p-values	AIC
Age at time of study	0.0451	-0.0010	p = 0.865	231.75
Age of onset	0.3735	-0.0304	p < 0.01	229.11
Age of arrival	1.0175	-0.0385	p < 0.01	213.81
Years in the U.S.	-0.1768	0.0070	p = 0.306	234.34
Education	-1.4439	0.3630	p < 0.003	166.28
Age of onset : Years in the U.S.	-0.6678	-0.0036	p < 0.01	221.25
Age of arrival: Years in the US	0.5093	-0.0006	p = 0.697	216.71
Age of onset : Age of arrival	-0.2919	-0.0050	P = 0.0973	208.57
Age of onset: Education	-2.0825	-0.0236	p = 0.2408	180.72
Age of arrival: Education	-0.4811	0.0059	p = 0.701	169.42
Years in the U.S.: Education	-2.3502	-0.0088	p = 0.4018	183.89

Table 14. Binomial logistic regression results for external factors predicting idiom accuracy

55 The regression results reveal that several external variables significantly predict idiom accuracy among L2 learners. A later age of onset is associated with decreased idiom performance ($p < 0.01$), suggesting that early exposure to L2 provides a lasting advantage in learning money idioms. Age of arrival in the U.S. also emerges as a strong predictor ($p < 0.01$), with later arrival linked to poorer performance, likely due to reduced opportunities for authentic, immersive exposure.

56 Years of residence in the U.S. alone is not a significant predictor of idiomatic performance ($p = 0.306$), indicating that mere duration of stay does not guarantee higher idiom proficiency. Education, however, stands out as a robust predictor: learners with higher education levels demonstrate significantly greater idiom recognition accuracy ($p < 0.003$), pointing to the importance of formal training and literacy-rich environments in the development of idiomatic competence.

57 Interaction models offer additional insights. The combination of age of onset and years in the U.S. yields a significant interaction ($p < 0.01$), indicating that early learners benefit more from extended residence than those who begin learning English later in life. However, the interaction between age of arrival and years in the U.S. is not statistically significant ($p = 0.697$). This suggests that living longer in the U.S.

does not compensate for a late arrival. The critical factor is when the exposure begins, not how long it lasts.

58 Finally, none of the interactions involving education––whether with age of onset, age of arrival, or years in the U.S.––reached statistical significance. This pattern indicates that education has its own distinct effect on idiom proficiency and does not change how factors like age of onset or age of arrival impact learning. Taken together, these findings underscore the primacy of early language exposure and educational attainment in developing money idiom recognition accuracy.

4. Discussion

59 This study investigated the relationship between L2 knowledge of money-related idioms and self-rated proficiency in English, focusing on whether idioms from a semantically homogenous domain could serve as an indicator of overall proficiency and its specific components (speaking, understanding, reading). Three additional research questions structured the investigation: (1) whether certain money idioms are acquired earlier than others and thus can differentiate proficiency levels; (2) whether idiom knowledge distinguishes native speakers from highly proficient L2 learners; and (3) how experiential variables influence idiom performance. All participants completed a 12-item idiom recognition task that tapped into their comprehension skills.

60 The first line of analysis compared idiom test performance by self-rated English proficiency levels. The median scores on the idiom test aligned well with participants' self-designated proficiency levels, revealing a clear downward trend in accuracy from Native-Like to Intermediate speakers. However, while the adjacent Native-Like and Advanced groups showed distinct median scores (7 vs. 5), both exhibited substantial internal variability. Statistical analyses confirmed that there were no significant differences between these two groups, despite their divergent central tendencies. Importantly, both the Native-Like and Advanced groups differed significantly from Native speakers on one end and from Intermediate learners on the other, indicating that while self-assessment provides a broad indication of

proficiency, it may not fully capture fine-grained differences in idiom knowledge between adjacent groups of advanced speakers.

61 The results of the Spearman rank correlation analysis further illuminate the connection between idiom scores and specific components of proficiency. A statistically significant correlation was found between test scores and self-rated speaking ability ($p < .0001$), suggesting that learners who perceive themselves as having strong oral production skills tend to perform better on the idiom recognition task. A weaker, yet still significant, correlation emerged for listening skills ($p < .05$), while reading ability did not reach significance. These findings align with Vanderniet's (2015) conclusion that idiom test performance––assessed via a comprehension task––correlates most strongly with speaking proficiency, reinforcing the close connection between oral competence and idiom mastery.

62 To investigate individual idiom patterns and how they vary across proficiency levels, we reclassified participants into idiom-based groups (G1, G2, G3). This reclassification yielded groups with greater internal consistency, as indicated by standard deviations comparable to those of the Native group. The validity of these idiom-based groupings was supported by statistical analyses: Welch's ANOVA revealed significant differences between groups ($p < 0.001$), and post-hoc Games-Howell tests confirmed robust pairwise contrasts – most notably between Native and G1 speakers, but also between adjacent levels (G1 vs. G2, G2 vs. G3). These results confirm the findings from the self-rated proficiency analysis, demonstrating that money idioms used in the test broadly function as indicators of non-native speaker status, regardless of the specific proficiency level within the L2 sample.

63 With respect to the question about idiom difficulty and acquisition, idiom-level analyses revealed a stable cluster of highly recognizable idioms––pretty penny, not buy it, won't break the bank, turn on a dime, and, to a slightly lesser extent, IOUs––across L2 groups. These idioms consistently ranked among the most accurately identified by all learner groups, suggesting they may function as threshold items that mark the transition from intermediate to advanced idiomatic competence. Of the four, only pretty penny showed partial lexical overlap with the learners' L1s through words like kopeyka or

grosh, which refer to small monetary units. The other three idioms had no direct equivalents in the participants' native languages, suggesting that their accurate recognition cannot be attributed to L1 transfer. This pattern points to the limited role of cross-linguistic influence in the acquisition of this idiom set and highlights the importance of other factors, such as input frequency, semantic transparency, and contextual salience. These findings may suggest L2 idiom learning––at all stages––may be shaped as strongly by perceptual and usage-based factors as by direct L1 influence (Wulff, 2019).

64 Frequency, familiarity, and phonological salience of individual lexical items within idioms likely play a critical role in their learnability and retention by L2 learners. Idioms like I don't buy it and *It* won't break the bank contain highly frequent and concrete lexical items, which may enhance their learnability through repeated exposure in the input. Likewise, the phonological salience and alliteration in pretty penny may boost memorability. The lexical item dime stands out as a culturally and idiomatically rich term in American English, appearing not only in turn on a dime but across a wide range of other idiomatic expressions and collocations. These include idiomatic phrases such as a dime a dozen, squeeze every dime out of someone, on the taxpayer's dime, and drop a dime. Other expressions like nickel-and-dime someone and worth every dime further boost the lexical salience of dime in everyday discourse. The prevalence of *dime* in figurative speech may thus enhance its learnability and reinforce its role as a marker of idiomatic proficiency in American English. While this interpretation remains speculative, it points to the need for a more detailed frequency-based analysis of idioms and their constituent lexical items in future work, particularly with respect to learner input. Such analyses could help clarify the relative contributions of frequency, transparency, and cultural salience in shaping idiom acquisition patterns.

65 Finally, regression modeling showed that external factors meaningfully influenced idiom performance. Age of onset of English learning and age of arrival in the U.S. were both significant predictors ($p < 0.01$), confirming that earlier exposure provides a lasting advantage in mastering this set of money idioms. Education level also emerged as a strong, independent predictor ($p < 0.003$), with higher levels of education associated with better idiom performance. In

contrast, years of residence in the U.S. was not a significant predictor on its own, suggesting that sheer duration of exposure does not ensure high test performance without considering the timing and quality of input.

66 These findings have implications for both language assessment and pedagogy. First, idiom accuracy emerges as a potentially valuable metric for distinguishing not only broad proficiency levels but also separate components of proficiency, especially speaking. Second, the gradation in idiom performance reinforces the notion that proficiency in idiomatic language develops incrementally. Rather than viewing idioms as uniformly "advanced" content, educators might consider introducing certain idioms earlier in instruction, especially those like won't break the bank or not buy it, which contain very common lexical items and which learners (at least learners of Slavic background) appear to acquire sooner. Conversely, more opaque or culturally dense idioms may require contextual scaffolding and explicit instruction in metaphorical reasoning to become accessible.

67 This study also illustrates the potential value of idiom-based assessments for curriculum design and learner placement (Gyllstad & Schmitt, 2019). An idiom recognition task could be implemented as a placement tool or progress diagnostic to differentiate between intermediate and more advanced learners, as well as between L2 learners and native speakers. It remains to be seen, however, how learners from other L1 backgrounds would perform on the task designed for this study. Future research could investigate whether L2 English speakers of typologically closer L1s demonstrate different patterns of idiom comprehension compared to learners from typologically more distant languages.

5. Conclusion

68 In conclusion, this study demonstrated that idiomatic language within a restricted semantic field––money-related expressions––has the potential to serve both as an indicator of L2 English proficiency and as a marker of non-native status. Self-reported overall proficiency aligned broadly with idiom performance, while speaking and listening skills showed a statistically significant correlation. Reclassifying participants into idiom-based proficiency groups yielded a

more internally consistent measure of learners' performance and revealed stage-like acquisition patterns. The emergence of a stable cluster of early-acquired idioms––pretty penny, not buy it, won't break the bank, turn on a dime––highlights the role of transparency, frequency, and contextual salience in shaping idiom knowledge, and, to a lesser degree, the role of cross-linguistic transfer.

69 Most studies on idioms rely on lists of semantically diverse expressions, which may obscure patterns tied to specific conceptual domains. Future L2 research could focus on idioms within targeted semantic fields such as emotions (e.g. *blow off steam*), food (e.g. *have a lot on one's plate*), or body-related expressions (e.g. *get cold feet*), among others. Our findings have implications for both assessment and instruction, suggesting that idiom accuracy can help distinguish not only among proficiency levels but also inform targeted pedagogical strategies. Ultimately, idiom learning proves to be a rich and revealing avenue for exploring and evaluating multiple dimensions of L2 knowledge.

BIBLIOGRAPHIE

Abel, B. (2003). English idioms in the first language and second language lexicon: A dual representation approach. Second Language Research, 19(4), 329–358. https://doi.org/10.1191/0267658303sr226oa

Aljabri, S. S. (2013). EFL students' judgments of English idiom familiarity and transparency. Journal of Language Teaching and Research, 4(4), 662–669. doi:10.4304/jltr.4.4.662-669

Bally, Ch. (1921). Traité de stylistique française. Heidelberg.

Bell, N. (Ed) (2017). Multiple perspectives on language play. De Gruyter Mouton.

Carrol, G., Conklin, K., & Gyllstad, H. (2016). Found in translation: The influence of the L1 on the reading of idioms in a L2. Studies in Second Language Acquisition, 38, 403–443. doi:10.1017/S0272263115000492

Carrol, G., & Conklin, K. (2017). Cross-language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'. Bilingualism: Language and Cognition, 20(2), 299–317. doi:10.1017/S1366728915000103

Cenoz, J., Hufeisen, B., & Jessner, U. (Eds) (2003). The multilingual lexicon. Kluwer Academic Publishers.

Charteris-Black, J. (2002). Second language figurative proficiency: A comparative study of Malay and

English. Applied Linguistics, 23(1), 104–133.

Cieślicka, A. B. (2006). Literal salience in online processing of idiomatic expressions by second language learners. Second Language Research, 22, 115–144. https://doi.org/10.1191/0267658306sr263oa

Cieślicka, 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. Journal of Psycholinguistic Research, 42, 475–503. doi: 10.1007/s10936-012-9232-4. PMID: 23161392

Cieślicka, A. B. (2015). Idiom acquisition and processing by second/foreign language learners. In R. R. Heredia & A. B. Cieślicka (Eds.), Bilingual figurative language processing (pp. 208–244). Cambridge University Press.

Gibbs, R. W. (1985). On the process of understanding idioms. Journal of Psycholinguistic Research, 14(5), 465–472. doi.org/10.1007/BF01666721

Guo, Y. (2014). Chinese EFL learners' comprehension of English idiom variants: Effects of variation type and proficiency level. Chinese Journal of Applied Linguistics, 37(4), 430–450.

Gyllstad, H., & Schmitt, N. (2019). Testing formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), Understanding formulaic language: A second language acquisition perspective (pp. 174–191). Routledge.

Irujo, S. (1986). A piece of cake: Learning and teaching idioms. ELT Journal, 40, 236–242. https://doi.org/10.1093/elt/40.3.236

Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. International Review of Applied Linguistics, 31, 205–219.

Isobe, Y. (2011). Representation and processing of formulaic sequences in L2 mental lexicon: How do Japanese EFL learners process multi-word expressions? JACET Kansai Journal, 13, 38–49.

Katz, A. N. (2024). A selective view of the last 50 years of the experimental psycholinguistics of idiom, metaphor, and irony: A commentary. Discourse Processes, 61(1–2), 90–102. https://doi.org/10.1080/0163853X.2024.2319898

Kellerman, E. (1986). An eye for an eye: Crosslinguistic constraints on the development of the L2 lexicon. In E. Kellerman & M. Sharwood Smith (Eds.), Crosslinguistic influence in second language acquisition (pp. 35–48). Pergamon Press.

Laufer, B. (2000). Avoidance of idioms in second language: The effect of L1-L2 degree of similarity. Studia Linguistica, 54, 186–196. https://doi.org/10.1111/1467-9582.00059

Lin, J.-T. (2016). Idioms in the bilingual mental lexicon. Doctoral dissertation. University of Hawai'i at Mānoa, USA.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. Journal of Speech, Language, and Hearing Research, 50(4), 940–967.

Mel'čuk, I. A. (2023). General phraseology: Theory and practice (1st ed.). John Benjamins.

Nacey, S. (2013). Metaphors in learner English. John Benjamins.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. Second Language Research, 27, 251–272. https://www.jstor.org/stable/43104668

Titone, D., Columbus, G., Whitford, V., Mercier, J., & Libben, M. (2015). Contrasting bilingual and monolingual idiom processing. In R. R. Heredia & A. B. Cieślicka (Eds.), Bilingual figurative language processing (pp. 171–207). Cambridge University Press.

Van Ginkel, W., & Dijkstra, T. (2020). The tug of war between an idiom's figurative and literal meaning: Evidence from native and bilingual speakers. Bilingualism: Language and Cognition, 23(1), 131–147.

Vanderniet, K. H. (2015). Idioms as a measure of proficiency. Master's thesis. Brigham Young University, USA.

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. Studies in Second Language Acquisition, 35, 451–482. doi:10.1017/S0272263113000107

Wu, X., Chen, B., & Huang, L. (2006). The processing of English idioms by Chinese EFL learners: Effects of idiom type and proficiency level. Foreign Language Teaching and Research, 38(3), 196–201.

Wulff, S. (2019). Acquisition of formulaic language from a usage-based perspective. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), Understanding formulaic language: A second language acquisition perspective (pp. 19–37). Routledge.

ANNEXE

Idiom test items

1. If I had a ___ for every time I've heard that said about me or my friends, I'd be driving a Maserati.

(a) cent

(b) penny

(c) nickel

(d) dime

(e) quarter

2. The problem is one we can solve, and it won't ___ the bank.

(a) empty

(b) clear

(c) hurt

(d) break

(e) hammer

3. I haven't spent one ___ cent in New York.

(a) sad

(b) sorry

(c) silver

(d) red

(e) green

4. You know, life really can turn on a ___.

(a) cent

(b) penny

(c) nickel

(d) dime

(e) quarter

5. Not sure what it's all about, but you can bet your ___ dollar I'm gonna try to get out of it.

(a) every

(b) hard-earned

(c) silly

(d) bottom

(e) top

6. Two pizzas and tip cost a ___ penny.

(a) paltry

(b) puny

(c) piney

(d) pretty

(e) pickled

7. In the sentence 'It is burning a hole in my pocket', the word *it* most likely means

(a) an unsolved problem

(b) a five-dollar bill

(c) a candy bar

(d) a piece of hot coal

(e) a utility bill

8. In the sentence 'However, the US is able to sell our IOUs to unsuspecting foreigners', the word *IOU* means

(a) bonds

(b) bitcoins

(c) bills

(d) checks

(e) fees

9. Based on the sentence ' "It's a very expensive job", says Jack, a 70-year-old sugar daddy', which of these has a meaning different from the other two?

(a) moneybags

(b) rainmaker

(c) sugar daddy

10. Which sentence has a meaning different from the other two?

(a) She makes dough.

(b) She brings home the bacon.

(c) She brings home veggies.

11. Which sentence has a meaning different from the other two?

(a) I am not into it.

(b) I don't buy it.

(c) I don't believe it.

12. Which of these is different in meaning from the other two?

(a) loaded

(b) in for a penny, in for a pound

(c) persistent

NOTES

1 In Titone, Columbus, Whitford, Mercier and Libben (2015, p. 173), *familiarity* is defined as "speakers' subjective impression of how often an idiom is encountered in its written or spoken form, regardless of whether the figurative meaning of the phrase is known" and *literal plausibility* is defined as "idiom's potential to have a literal interpretation" (e.g. *skate on thin ice*). Semantic transparency is referred to as *semantic decomposability*, which is defined as "how individual meanings of the idiom's component words relate to the figurative meaning of the phrase."

2 The label "congruent" refers to idioms that exhibit complete overlap between L1 and L2 in linguistic form and conceptual meaning: He bit the dust has a fully congruent counterpart in French––Il a mordu la poussière––which contains equivalent lexical items and conveys the same meaning.

3 One reviewer asks how the author ensured that participants did not use dictionaries or other external tools during the idiom test. To minimize this possibility, participants were explicitly instructed not to consult any external resources while completing the test. Although we did not implement technological controls (e.g. browser monitoring), the presence of numerous idiom-related errors and consistent accuracy gradient corresponding to self-reported proficiency levels provide indirect evidence that responses reflect genuine linguistic knowledge. Furthermore, participants had no incentive to artificially enhance their scores as they understood that

the study was purely for research purposes and that their responses were fully anonymized and depersonalized.

4 One reviewer noted that our idiom selection—based on two native speakers' responses—is subjective and potentially problematic, given evidence that many idioms are often infrequent in authentic language. We appreciate this well-taken point. We note that idiom selection procedures vary widely across studies, and how idioms are selected for L2 research could constitute a methodological paper in its own right. For example, Abel (2003) selected idioms from dictionaries (Longman Dictionary of English Idioms), while Titone et al. (2015, p. 182) selected idioms based on a structural characteristic: "we selected 120 English idioms having a pronoun-verb-x-noun structure" (without any further detail about the selection procedure itself). Idioms were then "coded for cross-linguistic overlap based on the judgments of two native speakers of Canadian French, supplemented by published and online French–English idiom dictionaries." In our study, idioms were not chosen based on estimated frequency, but rather on native-speaker familiarity: two native speakers independently listed money idioms they personally used or heard in everyday speech. Our selection procedure also incorporated a randomization component, which reduced researcher bias in item choice. Moreover, the Native group performed with high accuracy, suggesting that the idioms selected were indeed familiar and recognizable, thereby supporting their appropriateness as stimuli for investigating idiom recognition in L2.

5 The idiom in for a penny, in for a pound was not presented as a standalone sentence but was included as one of the multiple-choice options, based on its appearance as a full sentence in COCA.

6 One reviewer noted that the distractor item in (3) allows for more than one plausible answer. This was intentional. Some distractors were designed to be somewhat ambiguous to reduce pattern recognition and test-taking strategies. For example, while callow is synonymous with green (inexperienced), it may also be associated with callous due to phonological similarity. While some distractors permitted multiple interpretations, participants were instructed to choose the best answer. Only idiom items had one clearly correct response and were the focus of analysis; distractors served to create a more natural and cognitively engaging task.

7 One reviewer asks whether "some learners might have received explicit instruction about some of the selected idioms before the test." Participants were not informed about the nature of the test in advance, nor was any

explicit instruction on the idioms provided before the test. The instructions presented the task in general terms, stating: "In this survey (36 short questions), you will be asked to think about the meaning of various words and sentences. Some questions will ask you to fill in a blank by choosing a word from a multiple-choice menu. Other questions will ask you to decide which words or sentences are similar in meaning." Participants were instructed to choose only one best answer per question.

8 The item featuring sugar daddy was originally intended to be presented within the sentence *It's a very expensive job, says Jack, a 70-year-old sugar daddy*, drawn from COCA. However, this sentence was not displayed, and the survey presented only the multiple-choice options. As a result, this item was excluded from the analysis.

RÉSUMÉS

English

Idiomatic expressions can be challenging for L2 learners due to their metaphorical nature, cultural references, and resistance to compositional analysis. This study examines the role of idiomatic knowledge in proficiency assessment, focusing exclusively on English money idioms (e.g. *turn on a dime*, *pretty penny*). Given their rich metaphorical content, it is argued that these idioms provide a unique avenue for evaluating English proficiency and distinguishing between learners at different proficiency levels. The study analyzes 52 L1-Slavic/L2-English and 45 native English speakers, assessing their idiom knowledge via an online idiom recognition task. Results show that idiom accuracy rates broadly correlate with self-reported proficiency and especially oral skills (speaking and listening). It is also shown that regrouping L2 speakers based on idiom performance allows us to identify stage-like acquisition patterns. These results have implications for both language assessment and instruction, highlighting idiom accuracy as a means of refining proficiency classification and informing targeted pedagogical strategies.

Français

Les expressions idiomatiques peuvent être difficiles pour les apprenants de L2 en raison de leur nature métaphorique, de leurs références culturelles et de leur résistance à l'analyse compositionnelle. Cette étude examine le rôle des connaissances idiomatiques dans l'évaluation des compétences, en se concentrant exclusivement sur les expressions idiomatiques liées à l'argent en anglais (par exemple *turn on a dime*, *pretty penny*). Compte tenu de leur contenu fortement métaphorique, ces expressions sont considérées comme un moyen unique d'évaluer la maitrise de l'anglais et de distinguer les apprenants selon leur niveau de compétence. L'article étudie 52 locuteurs

natifs de L1-Slave/L2-Anglais et 45 locuteurs natifs de l'anglais et évalue leur connaissance des expressions idiomatiques à l'aide d'une tâche de reconnaissance d'expressions idiomatiques en ligne. Les résultats montrent que les taux de précision des expressions idiomatiques sont généralement corrélés aux compétences autodéclarées, en particulier les compétences orales (expression et compréhension). Ils démontrent également que le regroupement des locuteurs de L2 en fonction de leur maîtrise des expressions idiomatiques permet d'identifier certaines étapes dans l'acquisition d'une L2. Ces résultats ont des implications tant pour l'évaluation linguistique que pour l'enseignement des langues. Ils soulignent l'intérêt de la précision idiomatique pour affiner la classification des niveaux de compétence et développer des stratégies pédagogiques ciblées.

INDEX

Mots-clés
expressions idiomatiques, compétence idiomatique, compétence en langue seconde, évaluation linguistique

Keywords
idioms, idiomatic competence, second language proficiency, language assessment

AUTEUR

Elena Gavruseva
Linguistics Department, University of Iowaelena-gavruseva@uiowa.edu