Predictability and Familiarity Assessments for Greek Idiomatic Expressions: the Role of Reading Habits and Language Profiles

Anastasia Lada

Vrije Universiteit Brussel, Brussels, Belgium
Email: anastasia.Lada@vub.be
ORCID: https://orcid.org/0000-0003-1616-2891

Ifigeneia Dosi

Democritus University of Thrace, Komotini, Greece
Email: idosi@helit.duth.gr
ORCID: https://orcid.org/0000-0003-4911-2049

Philippe Paquier

Vrije Universiteit Brussel, Université Libre de Bruxelles, Brussels, Belgium
Universiteit Antwerpen, Antwerp, Belgium
Email: philippe.paquier@vub.be
ORCID: https://orcid.org/0000-0001-8565-7754

Christina Manouilidou

University of Ljubljana, Ljubljana, Slovenia
Email: christina.manouilidou@ff.uni-lj.si
ORCID: https://orcid.org/0000-0002-9523-6491

Stefanie Keulen

Vrije Universiteit Brussel, Brussels, Belgium
Email: stefanie.keulen@vub.be
ORCID: https://orcid.org/0000-0003-4582-5199

═══

Abstract:

This study examines the predictability and familiarity ratings of 199 Greek idiomatic expressions, building on the random sample used by Lada et al. (2024). The primary goals are to explore correlations between predictability, familiarity, and idiom dimensions such as decomposability, subjective frequency, and ambiguity, and to investigate the relationships between idiom familiarity, predictability, and participants' bilingual/multilingual profiles and reading habits. Sixty-three native Greek-speaking students at Democritus University of Thrace completed familiarity and predictability assessments based on a random selection of idioms from Vlaxopoulos (2007). Correlational analyses, aligned with Lada et al. (2024), show that subjective frequency is positively correlated with both ambiguity and decomposability. In addition, familiarity is weakly correlated with ambiguity, moderately correlated with decomposability, and strongly correlated with subjective frequency and predictability. Furthermore, predictability is weakly correlated with ambiguity, moderately correlated with subjective frequency and decomposability, but strongly correlated with familiarity. Logistic regression analyses reveal that the number of foreign languages spoken negatively predicts correct idiom completion in the

predictability task, with more languages associated with lower predictability scores. Mixed-effects linear models indicate that higher reading frequency is linked to lower familiarity ratings, whereas more books read is associated with higher familiarity. These findings provide novel insights into idiom comprehension among bilinguals, highlighting the influence of language profiles and reading habits on idiom familiarity and predictability. Limitations include the binary approach to predictability scoring and the lack of language-specific details. We hence suggest future studies consider typological factors and alternative results' interpretation for idiom predictability.

1. Introduction

In literal language processing, the meanings of individual lexical and phrasal components are combined to derive the overall meaning of an expression or a sentence. However, this compositional approach is insufficient for idiom processing. Idioms are expressions whose meaning cannot be derived by combining the meanings of their constituent words (Glucksberg, 1991). On the contrary, their comprehension depends on other factors including not only their complex linguistic characteristics, but also prior knowledge, contextual cues as well as grasping the speaker's intentions and recognizing cultural conventions (Nunberg et al., 1994). Therefore, idiom processing refers to the cognitive and linguistic mechanisms involved in recognizing and interpreting them. In the same vein, successful idiom processing might be facilitated by frequent reading (Cain et al., 2009) and/or speaking multiple languages (Wolter & Gyllstad, 2011). However, such factors are not thoroughly examined in the up-to-date literature, especially in those understudied languages such as Greek.

2. Literature Review

2.1 Idiom dimensions

Idioms vary across several linguistic dimensions that are thought to influence how they are processed in the brain (Lada et al., 2023). Some key dimensions are familiarity, frequency, ambiguity, decomposability, and predictability, all of which contribute to idiom linguistic complexity (Lada et al., 2024; Sprenger et al., 2019). Familiarity refers to the extent to which a person knows or recognizes an idiom. Frequency relates to how often an individual encounters an idiom in speech or writing. Ambiguity is concerned with whether an idiom can be interpreted both literally and figuratively (ambiguous idioms) or solely in a figurative sense (unambiguous idioms). Decomposability describes the degree to which the individual words within an idiom contribute to understanding its figurative meaning. Predictability, on the other hand, refers to how easily a missing word in an idiomatic expression can be guessed, with more predictable idioms being easier to complete or understand when partially given.

2.2 Familiarity and predictability as key dimensions

Among the various idiom dimensions affecting idiom processing, familiarity and predictability play a significant role. These linguistic dimensions underscore the cognitive complexity involved in idiom

processing. For instance, concerning familiarity, Schweigert (1986) demonstrated that in a self-paced reading task, participants took longer to read unfamiliar idioms compared to familiar ones. Later, Haibo and colleagues (2017) found that during a semantic judgment task paired with EEG, participants exhibited higher accuracy rates for familiar idioms, while event-related potential (ERP) data revealed an earlier N400 (mainly used to investigate semantic processing) latency for familiar idioms, indicating faster processing. Similarly, Libben and Titone (2008) conducted a series of online and offline experiments—including whole-sentence meaningfulness judgments, word-by-word self-paced reading, and word-by-word fixed-rate meaningfulness tasks—and consistently found that idiom familiarity facilitated comprehension across all tasks. According to the authors, this facilitative effect likely reflects how idiomatic strings are stored and accessed in memory, suggesting that familiar idioms may be partially retrieved from memory during comprehension.

Furthermore, studies have consistently shown that idiom predictability too plays a significant role in idiom processing. For example, Hubbard and colleagues (2023) examined how idiom predictability influences electrophysiological responses during language processing. In their study, participants completed an acceptability judgment task, revealing a significant link between cloze probability and N400 and P600 amplitudes specifically for critical words in idiomatic contexts. This finding suggests that the extent to which compositional analysis is involved in understanding figurative language may depend on the idiom's predictability with more predictable idioms being more holistically processed but still engaging compositional processing. Furthermore, the authors observed greater gamma activity for more predictable idioms, indicating faster processing, as these idioms are easier to anticipate and complete. Along the same lines, Libben and Titone (2008) showed that predictable idioms facilitate faster and more efficient processing.

While various idiom dimensions influence idiom processing, predictability and familiarity are particularly important because they capture two key aspects of idiom processing: how easily an idiom is retrieved from memory (familiarity) and how easily it is inferred compositionally (predictability). These two dimensions are often correlated (e.g., Titone & Connine, 1994b; Libben & Titone, 2008; Bulkes & Tanner, 2017), but their precise interaction remains unclear: does higher predictability always lead to higher familiarity, or are there cases where idioms are familiar but not predictable (or vice versa)? Studying their relationship can provide insights into whether idioms are predominantly processed holistically (as stored lexical units) and/or compositionally (via word-by-word analysis).

Furthermore, it is important to consider that idiom processing heavily depends on a network of interrelated dimensions, including decomposability, frequency, and ambiguity. For example, prior research has shown that predictability and decomposability are linked - more predictable idioms tend to be more decomposable (Libben & Titone, 2008; Bulkes & Tanner, 2017) - but this has not been systematically examined in Greek. Similarly, familiarity may be influenced by frequency. However, some idioms are highly familiar despite being low in frequency. By examining how predictability and familiarity interact with the rest of idiom dimensions, this study aims to examine the way idioms are processed and how their idiom dimensions affect their reliance on memory and/or compositional processes. In addition, given that Greek idioms have not been systematically examined in this way, this research provides an opportunity to expand our understanding of figurative language processing beyond well-studied languages like English. Currently, only one study (Lada et al., 2024) has examined

Greek idioms, assessing 400 expressions for (subjective) frequency, ambiguity, and decomposability. However, data on Greek idioms' predictability and familiarity, and their potential correlations, are still lacking. This study aims to address this gap, providing new insights into idiom processing in Greek.

2.3 Idiom processing and reading habits

Reading habits are defined here as the patterns that people exhibit when it comes to reading. This encompasses different aspects such as the frequency and duration of reading as well as the types of material people engage with (Riffo et al., 2024). Research has shown that reading habits affect the development of cognitive skills that can facilitate inference-making (Kendeou et al.,2014). Specifically, Kendeou and colleagues (2014) explain that "inferences allow readers to construct meaningful connections between text and relevant background knowledge" (p.11) while adding that readers who are weak in inference-making exhibit difficulties in the comprehension of even simple texts. This could be justified as frequent readers are often exposed to multiple types of contexts that can strengthen their ability of inference-making.

Along these lines, inference-making is important for the successful processing of idiomatic expressions. Idiom familiarity and predictability are linked to reading habits. According to Snow (2002), proficient readers are typically skilled in recognizing words automatically and have extensive background knowledge, both of which aid in effective reading and can also enhance idiom familiarity. Automatic word recognition, for instance, allows readers to quickly process language, supporting the understanding of complex expressions such as idioms. Cain and colleagues (2009) also emphasize the importance of contextual inference, which is crucial for both reading comprehension and idiom processing. As people read more frequently, they become better at using the surrounding text to process ambiguous expressions such as idioms. This skill is especially useful when interpreting unpredictable idioms whose meanings cannot be easily retracted and are more reliant on compositional processes. Nippold and Duthie (2003) further point out that understanding idioms is part of lexical development that begins in childhood and continues as individuals are exposed to more complex language. Thus, we would expect that frequent readers are more likely to become adept at inferring their meanings through context, reinforcing their familiarity with these expressions over time as well as facilitating idiom predictability.

For the purpose of this research, we will primarily focus on reading habits related to literature in either of the languages spoken by participants.

2.4 Idiom processing and bilingualism

In today's increasingly interconnected world, finding individuals who exclusively speak a single language, commonly referred to as "pure monolingual speakers," has become a challenging endeavor. Therefore, this study places its focus on "functional" bilingualism. Studies by Pliatsikas and colleagues (2017) have demonstrated that simultaneous bilinguals and sequential bilinguals, with frequent use of their second language (L2) show the same volumetric changes, functional changes and display resembling activity. They argue that actual "language use" may be a factor as important as "age of acquisition" in relation to participant selection for studies on bilingualism.

Idiom processing and by extension idiom familiarity have frequently been linked to a bilingual advantage. More specifically, research has suggested that bilinguals may have an advantage when it comes to processing language, largely due to their enhanced cognitive flexibility and the ability to draw on a broader range of linguistic resources (Bialystok et al., 2012). This advantage, however, extends beyond simple knowledge of words, as bilinguals may experience cross-language effects even in the processing of multi-word expressions like idioms. For example, Wolter and Gyllstad (2011) observed that Swedish-English bilinguals processed word pairs more quickly when the collocations were congruent in both Swedish and English compared to those that only existed in one language. This finding indicates that bilinguals may experience facilitation when processing idiomatic expressions as well, particularly when those expressions overlap across their languages. Moreover, Carrol and Conklin (2014) conducted a study in which English-native and Chinese-English bilinguals were presented with English idioms, their Chinese translations, and control phrases (i.e., literal phrases) in a lexical decision task. The results indicated that bilinguals were faster at recognizing translated idioms than control phrases, and their recognition speed mirrored that of native English speakers when processing English idioms. This suggests that bilinguals may have access to more idiomatic expressions than monolinguals, leading to faster processing and greater familiarity, particularly when idioms in both languages share similarities.

The assumption that bilinguals can be more familiar with idiomatic expressions is not unanimously supported by research. While bilinguals may enjoy cognitive advantages and have access to more linguistic resources, their familiarity with idiomatic expressions in their first language (L1) can be influenced by various factors, including language dominance, frequency of language use, and the context in which they encounter idiomatic expressions (Du et al., 2021). For instance, bilinguals might have more exposure to idioms in their L2, which could reduce their familiarity with idioms in their L1. As a result, bilinguals might not always be more familiar with idioms overall, especially if they use their L1 less frequently. This highlights the importance of considering the context in which bilinguals are exposed to idioms when evaluating their familiarity with these expressions.

Most research on bilingualism tends to focus on how L1 influences L2 processing, but less attention has been given to how knowledge of an L2 might affect processing in L1, particularly in the context of idioms. There is evidence that even a less dominant L2 can influence how L1 is processed. For example, van Hell and Dijkstra (2002) showed that bilinguals responded faster in lexical decision tasks to L1 words that had cognates in their L2, suggesting that processing of L1 words can be facilitated by knowledge of L2. This phenomenon is consistent with the idea that bilinguals' language systems are interconnected, and that processing in one language can influence the other. In a similar vein, Du and colleagues (2021) argue that bilingual memory operates in a parallel, non-selective way, where the activation of one language can influence the processing of the other, even when the task is conducted entirely in L1. Their study, which involved English-Chinese bilinguals, found that L1 binomials (fixed pairs of words) that were congruent with their L2 counterparts showed faster lexical decision times, providing evidence that L2 processing can affect L1 even when the task is strictly in L1.

These cross-language effects have important implications for how bilinguals process idiomatic expressions. Although bilinguals' exposure to idioms in both languages might expand their idiomatic knowledge, providing them with a larger pool of idioms to draw from, which could lead to increased

familiarity with idioms overall, especially if there is overlap between idioms in the two languages, this advantage may come with a potential cost. If L2 idioms are activated during L1 processing, it could introduce interference, particularly if there are similar or congruent idioms across the two languages. This could complicate the task of interpreting idioms, as bilinguals might be more prone to confuse idiomatic expressions from both languages. Therefore, bilinguals' knowledge of idioms might not always result in better performance, particularly if there is a significant overlap between idioms in the two languages or if their L1 is underused compared to their L2. In sum, while bilinguals may have access to a broader range of idioms, this could lead to both advantages and challenges in idiom processing, also affecting their idiom familiarity and potentially predictability.

The relationships between idiom familiarity, reading habits, and bilingualism can also extend to idiom predictability, which is frequently correlated with idiomatic familiarity in various studies (e.g., Bulkes & Tanner, 2017; Tabossi et al., 2011). Notably, predictability plays a key role in idiom recognition, as it often requires the correct completion of an idiomatic string, making it another indicator of idiom processing and comprehension. This study's second aim is to explore the potential associations between reading habits, bilingualism (and multilingualism), and both idiom familiarity and predictability, offering a comprehensive view of how these factors might interact in processing Greek idiomatic expressions.

3. The Present Study

This study analyzed predictability and familiarity ratings for the same 199 Greek idiomatic expressions assessed by Lada et al. (2024) for their random sample. The main goal was to explore potential correlations between predictability, familiarity, and the other idiom dimensions tested in the earlier study. Additionally, this study aimed to examine the relationship between reading habits, bilingualism (and multilingualism) and idiom predictability and familiarity. Here, it is important to note that there are no studies, to the best of our knowledge, investigating the effects of predictability and familiarity in bilingualism. In this study, all participants spoke at least one foreign language and therefore, they are not considered as "pure monolinguals". To this end, this study seeks to answer the following research questions.

1) To what extent do predictability and familiarity correlate with each other in the dataset of Greek idioms?

2) How do predictability and familiarity correlate with other idiom dimensions such as (subjective) frequency, ambiguity and decomposability?

3) To what extent are predictability and familiarity predicted by participants' bilingual (and multilingual) profile, as well as their reading habits?

4. Methods

4.1 Participants

In this study, the authors recruited the same participants as in the study by Lada et al. (2024), specifically, sixty-three healthy, native Greek speakers studying Greek Philology at Democritus University of Thrace in Komotini, Greece. To gain further insights into participants' demographic profiles, they were asked to provide some personal information before beginning the questionnaire on idiom predictability. To gain further insights into their demographic profiles, participants were asked to provide some personal information before beginning the questionnaire on idiom predictability, including their educational level, number of foreign languages spoken (previously collected by Lada et al., 2024), the number of languages they use daily, their frequency of reading literature, and the number of books they read annually. These questions regarding reading habits were intended to explore potential correlations between participants' engagement with literature and their ability to predict idiomatic expressions, as well as to assess their general familiarity with idioms. Finally, it is important to note that all participants spoke at least one foreign language but not all of them are considered functional bilinguals since only a few use more than one language for their everyday communication needs.

4.2 Materials and design

For the assessment of predictability and familiarity, 200 idioms were randomly chosen from the dictionary of Vlaxopoulos (2007). The list of the idioms is identical to the one used by Lada et al. (2024) in their random sample and was used for both questionnaires created for this study. Specifically, every second or third idiom from each page of the Greek dictionary was selected until they reached a total of 200. However, in the predictability task, one idiomatic expression was not correctly presented in the questionnaire and therefore it was excluded from the analysis. Therefore, only 199 idioms are considered for this study. Similarly to Lada et al. (2024), we constructed two questionnaires aiming to test idiom predictability and familiarity following the research design by Libben and Titone (2008) and Titone and Connine (1994).

The idiom predictability questionnaire presented participants with a cloze task, where they were asked to fill in the missing word for each idiomatic expression. Specifically, the missing word was the final constituent of each idiom, selected because it significantly contributes to the idiom's figurative meaning. If the last word served mainly a syntactic function without adding lexical or conceptual weight, the penultimate word was removed instead. Next, the questionnaire on idiom familiarity asked participants to rate how well they know each idiomatic expression using a six-point Likert scale.

1) Predictability Assessment

   All participants completed a cloze task consisting of 199 idiomatic expressions embedded in carrier sentences with neutral contexts. They were instructed to fill in the gap with the word that best completed each idiomatic expression. To aid their responses, the meaning of each idiom was provided in parentheses immediately following the carrier sentence. If participants could not determine the missing word, they were asked

to enter an "X" or "-" in the gap. Below, the instructions translated into English are provided.

Instructions

This questionnaire includes different idioms in Greek. Each idiom is placed between single quotes. All the following idioms have a missing word. For each idiom, you are asked to fill in the word that best completes the idiomatic expression. Write the word you think is missing in the blank provided. To help you, we suggest you consider the meaning of each idiom between brackets. In case you find it difficult to find the missing word, write a hyphen (-) or an X in the blank.

2) Familiarity Assessment

All participants completed a questionnaire with the same 199 idiomatic expressions. However, the idioms were not presented within carrier sentences to avoid context biasing their judgments. They were instructed to decide on how well they knew the idioms presented to them. There were no idiom definitions provided. Participants had to decide using a rating scale ranging from (0), defined as "I do not know its meaning at all", to (5), defined as "I know its meaning very well". Below, the instructions translated into English are provided.

Instructions

This questionnaire includes different idioms in Greek. All of them have a metaphorical meaning. For the following idioms, you must decide whether you know their meaning. Your scores will be on a scale from 0 to 5 where number 0 means that an idiom is completely unknown and you do not know what it means at all, while number 5 means that an idiom is very well known, and you know very well what it means. The intermediate values on the scale should indicate your judgment about how well you know each idiom. Use the full scale to indicate your judgments.

4.3 Procedure

The procedure replicated that of Lada et al. (2024), as this study complements their work by providing data on two distinct dimensions of idiom comprehension. Consequently, all sixty-three participants completed both assessments. As outlined by Lada et al. (2024), a within-subjects design was chosen to gather normative data, given that these ratings are not independent measures but instead reflect participants' deeper perceptions of idiomatic expressions. Participants were tested in two different sessions, each focusing on the assessment of an idiom dimension. First, participants were asked to complete the questionnaire on idiom predictability. This session lasted approximately 40 minutes. Then, to avoid participant fatigue, participants were required to complete the second questionnaire on a different day. The participants completed the first questionnaire on predictability before completing any other questionnaire, including those administered in the context of the study by Lada et al. (2024). This is important since the questionnaire on predictability asks participants to fill one missing word in the idiomatic expression and therefore, earlier contact with any other questionnaire

would provide them with the answers. Both assessments were administered online, through Google Forms and they were automatically randomized for stimulus order.

5. Results

5.1 Participant profiles

The study included healthy participants, aged 19-39 years (M = 22.2, SD = 2.94), 58 females, 3 males, and 2 individuals who preferred not to disclose their gender. All participants were native Greek speakers, free of any neurological disorders, and had received tertiary education: 44.4% had 1-3 years of higher education, 42.9% had 4-6 years, and 12.7% had more than 6 years. English was their second language, with 44.4% speaking one foreign language, 41.3% speaking two, and 14.3% speaking three, while none could communicate in more than three languages. This demographic information aligns with data provided in the study by Lada et al. (2024). In terms of daily language use, 47.7% used one language, another 47.7% used two languages, and 4.6% used more than two languages, indicating that over half of the participants were functional bilinguals (i.e., actively using more than one language for their everyday communication needs). Regarding reading habits, 3.1% reported not reading literature at all, 18.5% read occasionally during the week, 23.1% read monthly, and 53.8% read occasionally throughout the year, with one participant not responding. Additionally, 4.6% reported reading zero books annually, 78.5% read 1-5 books per year, 12.3% read 6-10 books, and 4.6% read more than 10 books per year. The questions about participants' reading habits focused on literature reading but did not differentiate between recreational reading and reading within educational contexts. Consequently, it's worth noting that all participants were university students and therefore, on average, frequent readers, likely exhibiting higher-than-average reading habits compared to the general population.

5.2 Database

In Appendix A, all 199 idioms are presented with participants' answers on familiarity as well as their responses in the predictability cloze task for 199 idioms.

5.3 Reliability of familiarity ratings

To assess the internal reliability of familiarity ratings, the intraclass correlation coefficient (ICC) was calculated following Lada et al. (2024) and similar to Hubers et al.' (2019). The internal reliability was calculated with the parameters "two-way mixed" and "absolute agreement". The ICC calculations showed that the mean ICC was 0.96 with a 95% confidence interval ranging from 0.95 to 0.97 (F = 45.353, p < 0.001).

5.4 Descriptive statistics

The average rating, standard deviation, range, minimum and maximum values, and total number of idioms analyzed for predictability and familiarity are presented in Table 1 below. Figure 1 shows how the ratings of familiarity and predictability are distributed across the scales similar to Lada et al.' (2024). Last, Table 2 presents examples of high and low predictable and high and low familiar idioms.

Table 1. Descriptive Statistics for Predictability and Familiarity of the Greek Idioms

	Familiarity	Predictability
Average rating	4.21	0.51
Std. Deviation	0.84	0.36
Range	4.03	1.00
Minimum	0.97	0.00
Maximum	5.00	1.00
Total No. of Idioms	200	199

Mean = 4.22
Std. Dev. = .845
N = 200
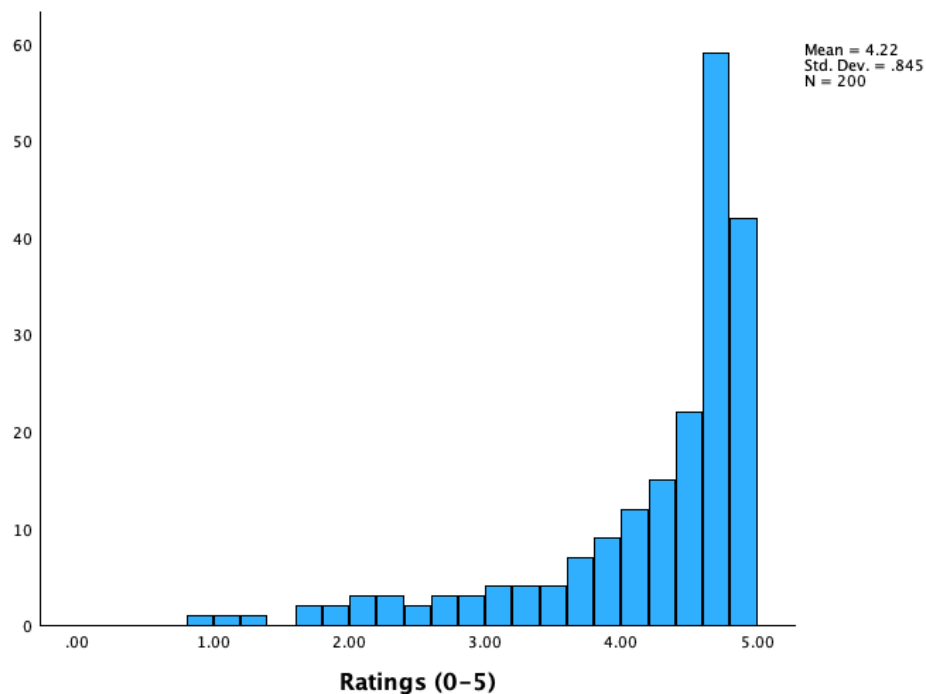
Figure 1. Histogram Plots Showing the Distribution of All Greek Idioms for the Ratings of Familiarity (First Histogram) and Predictability (Second Histogram). Indications on a Likert Scale of 0-5 for Familiarity and 0-1 for Predictability
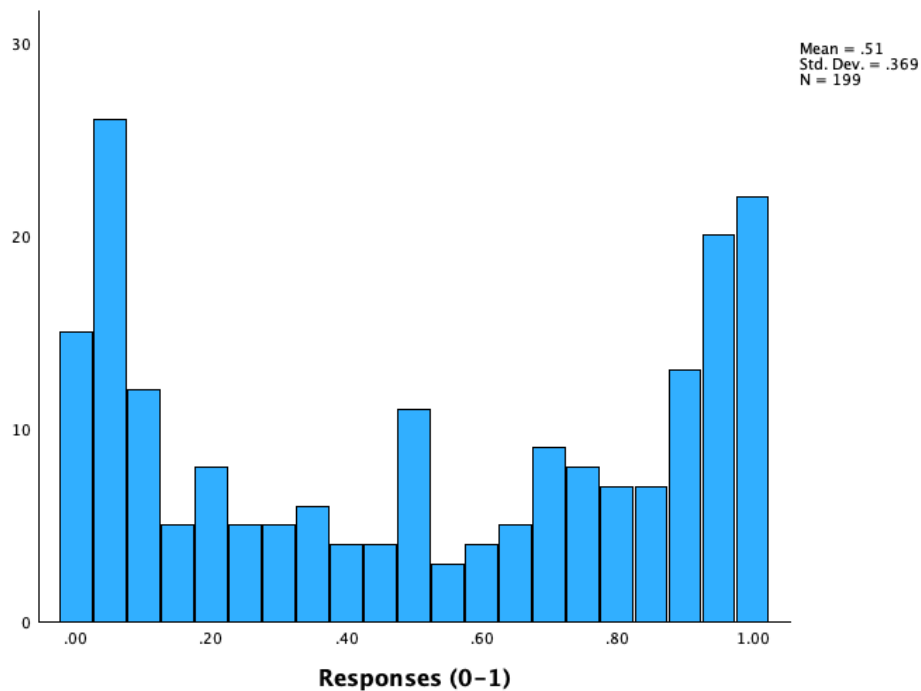
Table 2. Examples of Idioms Rated as High or Low for Familiarity and Predictability. Each Example Shows the Greek Idiom and its Figurative Meaning in Parentheses, Followed by the English Translation

	Low	Av. Rate	High	Av. Rate
Familiarity	Χάρις στην κεραμιδιά μου (=χάρις του έρωτά μου) Favoring my tile (=because of my love for someone)	(0.97)	Κρατάω το φανάρι (= διευκολύνω την ερωτική συνεύρεση) I keep the lantern (= I facilitate the sexual encounter)	(5.00)
	Κάτι παίρνει τα ύψη (=ανεβαίνει ψηλά) Something takes the heights (=Something goes high in the sky)	(1.98)	Τα χαλάσαμε (=διαφωνήσαμε) We broke them down (=we disagreed)	(4.72)
Predictability	Κάνω το νερό μου (= ουρώ) I do my water (=I pee)	(0.03)	Με παίρνει ο ύπνος (= αποκοιμούμαι) The sleep takes me (=I fall asleep)	(1.00)
	Βαστάω χαρακτήρα (= μένω πιστός) I keep character (=I remain faithful)	(0.02)	Έχω σώας τας φρένας (=είμαι λογικός) I have safe brakes (=I am in my right mind/ I act logically)	(0.97)

5.5 Correlational analyses

For this study we have included the normative data collected by Lada et al. (2024) in their random sample for subjective frequency, ambiguity and decomposability. However, since predictability responses were available for 199 idiomatic expressions, we excluded that single idiomatic expression from all the rest of the datasets. More specifically, the correlational analyses showed data almost identical to Lada et al.' (2024) results regarding subjective frequency, ambiguity and decomposability. More specifically, subjective frequency was positively correlated with both ambiguity (r = .288, df = 199, p < .01) and decomposability (r = .670, df = 199, p < .01). Therefore, the degree of an idiom's subjective frequency is weakly related to the degree of its ambiguity and strongly to its degree of decomposability. In addition, decomposability was positively, and moderately correlated with ambiguity (r = .355, df = 199, p < .01).

Next, considering idiom predictability and familiarity positive correlations were found with all the rest of the idiom dimensions. Specifically, familiarity was weakly correlated with ambiguity (r = .178, df = 199, p < .01), moderately correlated with decomposability (r = .669, df = 199, p < .01), and strongly correlated with subjective frequency (r = .869, df = 199, p < .01) and predictability (r = .701, df = 199, p < .01). Next, predictability was weakly correlated with ambiguity (r = .267, df = 199, p < .01), moderately correlated with subjective frequency (r = .697, df = 199, p < .01), and decomposability (r = .500, df = 199, p < .01), while it was strongly correlated with familiarity (r = .701, df = 199, p < .01). Table 3, below, presents the correlations matrix.

Table 3. Correlations Matrix for all Idiom Dimensions

IDIOM DIMENSION	1	2	3	4	5
1. Subjective Frequency					
2. Ambiguity	.288**				
3. Decomposability	.670**	.355**			
4. Familiarity	.869**	.178**	.669**		
5. Predictability	.697**	.267**	.500**	.701**	

** Correlation is significant at the 0.01 level (2-tailed); N=199.

5.6 Effects of participant language profile and reading habits on predictability

To investigate predictability's link to participants' language profiles as well as their reading habits, we conducted binary logistic regression. The results showed that only the languages spoken by participants act as significant predictors for correctly completing the idiomatic expressions. More specifically, the logistic regression with the number of foreign languages spoken by participants as the independent variable showed that it significantly predicted the responses in predictability task. Interestingly, participants who spoke more foreign languages were less likely to respond correctly in idiom predictability. The model showed that with each additional foreign language spoken, the

likelihood of a correct response in predictability decreased by approximately 15%, as shown by the odds ratio (B= -.159, Exp(B)=0.853). The model was a moderate fit to the data as indicated by Hosmer and Lemeshow test (p=.010). Additionally, the model accurately classified 52.9% of the cases. The regression analyses showed no significant results for the number of languages participants used daily, the frequency of their literature reading or the number of books they read.

5.7 Effects of participant language profile and reading habits on familiarity

As a next step, we conducted mixed-effects linear models to examine whether the abovementioned parameters predicted participants' ratings on familiarity. Only participants' reading habits were significant predictors of the familiarity ratings. First, a mixed-effects linear model was conducted to examine whether the frequency of reading (ranging from 0 = never to 4 = daily) predicts familiarity ratings on a scale from 0 to 5. The fixed effect of frequency of reading was statistically significant, with an estimate of -0.240 (SE = 0.073, p = 0.001), indicating that higher frequency of reading was associated with lower familiarity ratings. Regarding the random effects, the variance of the random intercept was 0.223 (SE = 0.043), suggesting that there is notable variability in familiarity ratings between participants. The intraclass correlation (ICC) for the random intercept was 0.278 (adjusted) and 0.273 (conditional), indicating that a substantial proportion of the variability in familiarity ratings is due to between-participant differences. The model fit was assessed using the Akaike Information Criterion (AIC = 43342.1346) and the Bayesian Information Criterion (BIC = 43357.0072), both of which suggest an adequate fit. The significance of the fixed effect of frequency of reading was further assessed using Type III fixed effects, which revealed a statistically significant effect on familiarity ratings, $F (1, 107.606) = 10.706$, p = 0.001.

Then, a mixed-effects linear model was conducted to examine whether the number of books read (ranging from 0 = none to 3 = more than 10 books per year) predicts familiarity ratings on a scale from 0 to 5. The fixed effect of the number of books read was statistically significant, with an estimate of 0.156 (SE = 0.065, p = 0.044), indicating that a higher number of books read was associated with higher familiarity ratings. Regarding the random effects, the variance of the random intercept was 0.604 (SE = 0.252), suggesting that there is considerable variability in familiarity ratings between participants. The intraclass correlation (ICC) for the random intercept was 0.125 (adjusted) and 0.125 (conditional), indicating that a moderate proportion of the variability in familiarity ratings is due to between-participant differences. Model fit was assessed using the Akaike Information Criterion (AIC = 43308.7827) and the Bayesian Information Criterion (BIC = 43338.5278), both of which suggest an adequate fit. The significance of the fixed effect of the number of books read was further assessed using Type III fixed effects, which revealed a statistically significant effect on familiarity ratings, $F (1, 7.890) = 5.704$, p = 0.044.

6. Discussion

This study has provided complementary normative data for the random sample of 199 Greek idioms in Lada et al. (2024). Particularly, this study attempts to provide normative data on predictability and familiarity, investigate potential correlations between them as well as subjective frequency, ambiguity

and decomposability. In addition, since participants employed for this study could speak at least one foreign language, the study provides results concerning the effects of bilingualism or multilingualism as well as participants' reading habits on idiom predictability and familiarity.

First, concerning the relationship between familiarity and predictability, spearman's correlation coefficients showed that there is a significant strong positive correlation between the two dimensions in Greek idioms. This entails that for the Greek idioms analyzed, better-known idioms were also easier to complete in the cloze task. This agrees with other studies providing normative data showing significant strong correlations between predictability and familiarity (Bonin et al., 2013; Libben & Titone, 2008; Tabossi et al., 2011; Titone & Connine, 1994). Nevertheless, the results showed that this relationship is not arbitrary and there are instances of highly familiar idioms that are not very predictable. For example, the idiom "βγαίνω γελασμένος" (= I am deceived) was rated as a highly familiar idiom with a mean score of 4.34 on a Likert scale ranging from 0 to 5 where 5 corresponded to highly familiar and 0 to completely unfamiliar. However, the idiom was totally unpredictable with a score of 0.0 indicating that none of the participants was able to correctly complete the idiomatic string. This is very indicative of the dynamic and complex nature of idiomatic expressions (Titone and Libben, 2014).

Considering the relationship between idiom familiarity and other idiomatic dimensions, we found moderate to strong positive correlations with idiom decomposability and subjective frequency, and a weak positive correlation with ambiguity. This indicates that more familiar idioms are often encountered more frequently and are more decomposable (Tabossi et al., 2011; Nordmann et al., 2014; Gavilán et al., 2021). The connection between familiarity and frequency is intuitive, as idioms that are encountered more often are likely to be better acquired. Similarly, decomposable idioms may be more familiar because their constituent words facilitate easier processing, which in turn supports their acquisition. Lada et al. (2024) explained that when someone encounters a familiar idiom, they cannot inhibit their knowledge and assess an idiom's decomposability (Keysar & Bly, 1995). Last, the weak correlation between familiarity and ambiguity may suggest that idioms with both literal and figurative interpretations are more familiar. Nordmann and colleagues (2014) support that it is possible that when someone encounters an idiom that is well known, they lose their ability to inhibit their knowledge and assess an idiom's ambiguity very similar to what is happening between familiarity and decomposability. One possible explanation is that such idioms, with their dual meanings, might be encountered more frequently, leading to higher familiarity. Gibbs (1980, 1986) reported this as a memory advantage explaining that since ambiguous idioms need to be processed dually -literally and figuratively- they can be more easily recalled. Since they can be interpreted in multiple ways, they may appear more often in different contexts, reinforcing their recognition. However, given the weak nature of this correlation, it is difficult to draw definitive conclusions, and further research is needed to clarify this relationship.

Considering the relationship between idiom predictability and other idiomatic dimensions, we found moderate to strong correlations with idiom decomposability and subjective frequency, and a weak correlation with ambiguity, much like the pattern observed with idiom familiarity. These results suggest that more predictable idioms tend to be both more frequent and more decomposable. One possible explanation is that idioms encountered more often in everyday discourse are easier to predict when a

constituent word is missing (Libben & Titone, 2008). In a similar vein, idioms whose constituent words are closely linked to their figurative meanings are processed more efficiently, aiding their predictability and leading to more successful completion of the idiomatic string. Additionally, the relationship between predictability and ambiguity shows that more predictable idioms tend to be more ambiguous. This can be explained by the fact that ambiguous idioms are often more widely used and exposed in daily communication, making them more predictable despite their multiple meanings.

We also investigated whether bilingualism (and multilingualism) and participants' reading habits were predictors for their familiarity ratings and their successful responses in idiom predictability in a cloze task. First, the analyses showed that participants who spoke more foreign languages were less likely to correctly complete the idiomatic string in the cloze task assessing idiom predictability. This is in line with the studies showing that there is a cross-linguistic interference in bilingual speakers (Du et al., 2021). Bilinguals or multilinguals might experience interference from their foreign languages making it harder to find the correct missing word in the idiomatic string. At this point, we need to emphasize that in some idiomatic expressions participants provided responses that were semantically related to the correct word. Such partially incorrect responses might be linked to bilingual or multilingual processing showing that the meaning of the idiom is available and accessed but interference from second languages leads to cognitive overload and incorrect responses. Another explanation could be that people who speak less foreign languages have more exposure to their L1, in media and everyday discourse, leading to higher predictability of idioms in L1.

Second, the statistical analyses revealed that idiom familiarity was significantly predicted only by participants' reading habits. Notably, higher reading frequency was associated with lower familiarity ratings. This finding contradicts studies suggesting that frequent readers, who often encounter idioms in written language, tend to become adept at inferring idiomatic meanings through context, thus reinforcing familiarity over time (Snow, 2002). In contrast, in our study, participants who read more frequently reported lower familiarity with idioms. One possible explanation may relate to the type of reading participants engaged in. If they primarily read academic texts or materials with a refined, sophisticated writing style, this could account for lower familiarity with idioms, which are generally colloquial expressions encountered more often in casual, everyday language. This difference in exposure context could limit familiarity with idioms typically used in daily conversation rather than formal written materials.

Finally, the number of books participants read annually was linked to higher familiarity ratings. While this may seem counterintuitive given the findings on reading frequency, a possible explanation is that participants who read more books per year likely engage in longer, more immersive reading sessions. This type of engagement could be associated with better retention of idioms, in contrast to those who read frequently but in shorter sessions. However, these divergent results regarding reading habits might reflect differences in the genres participants read, their reading styles, and the depth of their engagement with the material. Additional information on participants' specific reading habits could offer further insights into these patterns.

7. Limitations and Future Directions

This study comes with certain limitations. First, the questionnaire designed to assess idiom predictability involved removing a single, key constituent word from each idiomatic expression (mostly at the ultimate or penultimate position), and only content words were removed. However, some idioms may contain multiple content words, each with varying significance for completing the idiom. Future research could refine the predictability assessment by accounting for the differing contributions of multiple content words in idioms. Additionally, the analysis of participant responses in the predictability task was conducted in a binary manner, thus excluding other potentially insightful responses. For instance, some participants gave semantically related responses that were incorrect for the target word but still demonstrated relevant understanding, while others provided responses that completed a different idiomatic expression correctly. Future studies could enhance predictability analysis by ranking answers based on the semantic similarity between the intended word and alternative responses. This is especially interesting in the case of Greek idioms where an idiom's constituent word could be replaced with another, and an alternative idiom would come up that would still be semantically and pragmatically equal (Mazi, 2014).

Furthermore, although participants reported their language profiles, there was no detail regarding the specific foreign languages they spoke. Typologically distant languages may influence language processing differently, which could impact idiom familiarity and/or recognizability. Additionally, factors like proficiency and language dominance, along with the specific genres or types of reading participants engaged in, could be considered in future studies. Similarly, the participants in this study were all Philology students at Democritus University of Thrace, meaning their reading habits may not align with average measurements. Finally, the inter-contradictory findings regarding reading habits and familiarity could be better understood with additional questions addressing the genres participants read, as well as their reading purposes and depth of engagement with the material.

8. Conclusion

This study provided complementary data on predictability measurements and familiarity ratings for 199 Greek idioms, which consisted of the random sample in the study by Lada et al. (2024). Analyses of the idiom dimensions revealed a strong positive correlation between idiom predictability and familiarity, with both dimensions showing moderate positive correlations with decomposability and subjective frequency, and a weaker positive correlation with ambiguity. Furthermore, the statistical analyses demonstrated that participants' reading habits significantly predicted familiarity ratings, while the number of foreign languages spoken was a significant predictor of idiom predictability.

Declarations and Acknowledgment:

References

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences, 16*(4), 240-250.

Bonin, P., Méot, A., & Bugaiska, A. (2013). Norms and comprehension times for 305 French idiomatic expressions. *Behavior Research Methods, 45*, 1259-1271.

Bulkes, N. Z., & Tanner, D. (2017). "Going to town": Large-scale norming and statistical analysis of 870 American English idioms. *Behavior Research Methods, 49*, 772-783.

Cain, K., Towse, A. S., & Knight, R. S. (2009). The development of idiom comprehension: An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology, 102*(3), 280-298.

Carrol, G., & Conklin, K. (2014). Getting your wires crossed: Evidence for fast processing of L1 idioms in an L2. *Bilingualism: Language and Cognition, 17*(4), 784–797.

Du, L., Elgort, I., & Siyanova-Chanturia, A. (2021). Cross-language influences in the processing of multiword expressions: From a first language to second and back. *Frontiers in Psychology, 12*, 666520.

Gavilán, J.M., Haro, J., Hinojosa, J.A., Fraga, I. & Ferré, P. (2021). Psycholinguistic and affective norms for 1,252 Spanish idiomatic expressions. *PloS One, 16*(7).

Geeraert, K., Baayen, R. H., & Newman, J. (2017). Understanding idiomatic variation: A usage-based approach. *Journal of Memory and Language, 94*, 291-311.

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & cognition, 8*(2), 149-156.

Gibbs Jr, R. W. (1986). Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse processes, 9*(1), 17-30.

Glucksberg, S. (1991). Beyond literal meanings: The psychology of allusion. *Psychological science, 2*(3), 146-152.

Haibo, S., Yan, W., & Wenyu, L. (2017). Familiarity in Chinese and English idiom comprehension: An event related potential study from Chinese English learners. *Chinese Journal of Applied Linguistics, 40*(1), 74-92.

Hubbard, R., Bulkes, N., & Lai, V. T. (2023). Predictability and decomposability separately contribute to compositional processing of idiomatic language. *Psychophysiology, 60*(8), e14269.

Hubers, F., Cucchiarini, C., Strik, H., & Dijkstra, T. (2019). Normative data of Dutch idiomatic expressions: Subjective judgments you can bank on. *Frontiers in Psychology, 10*, 1075.

Kendeou, P., Van Den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning disabilities research & practice*, *29*(1), 10-16.

Lada, A., Paquier, P., Manouilidou, C., & Keulen, S. (2023). A systematic review: Idiom comprehension in aphasia: The effects of stimuli and task type. *Journal of Neurolinguistics, 65*, 101115.

Lada, A., Paquier, P., Dosi, I., Manouilidou, C., Sprenger, S., & Keulen, S. (2024). Four hundred Greek idiomatic expressions: Ratings for subjective frequency, ambiguity, and decomposability. *Behavior Research Methods*, 1-15.

Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition, 36*, 1103-1121.

Mazi, V. D. (2014). *The use of journalistic texts in the teaching of Greek as a foreign language: the case of privatizations* (No. GRI-2014-13065). UK: Aristotle University of Thessaloniki.

Nippold, M. A., & Duthie, J. K. (2003). Mental imagery and idiom comprehension: A developmental study. *Journal of Speech, Language, and Hearing Research, 46*(4), 788-799.

Nordmann, E., Cleland, A. A., & Bull, R. (2014). Familiarity breeds dissent: Reliability analyses for British-English idioms on measures of familiarity, meaning, literality, and decomposability. Acta Psychologica.

Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language, 70*(3), 491-538.

Pliatsikas, C., DeLuca, V., Moschopoulou, E., & Saddy, J. D. (2017). Immersive bilingualism reshapes the core of the brain. *Brain Structure and Function, 222*(4), 1785-1795.

Riffo, B., Rojas, C., Helo, A., Véliz, M., Urzúa, P., Gutierrez, G., & Guerra, E. (2024). Reading Comprehension in Older Adults—Effects of Age, Educational Level, and Reading Habits. *Journal of Intelligence, 13*(1), 4.

Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research, 15*, 33-45.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Rand Corporation. Santa Monica, CA.

Sprenger, S. A., la Roi, A., & Van Rij, J. (2019). The development of idiom knowledge across the lifespan. *Frontiers in Communication, 4*, 29.

Tabossi, P., Arduino, L., & Fanari, R. (2011). Descriptive norms for 245 Italian idiomatic expressions. *Behavior Research Methods, 43*, 110-123.

Titone, D., & Libben, M. (2014). Time-dependent effects of decomposability, familiarity and literal plausibility on idiom priming: A cross-modal priming investigation. *The Mental Lexicon, 9*(3), 473-496.

Titone, D. A., & Connine, C. M. (1994). Comprehension of idiomatic expressions: effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1126.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*(4), 780-789.

Vlaxopoulos, S. (2007). *Dictionary of idioms in Modern Greek* (In Greek). Kleidarithmos.

Wolter, B., & Gyllstad, H. (2011). Examining the role of L1 lexical transfer on L2 comprehension: A Swedish–English bilingual perspective. *Language Learning, 61*(1), 69-100.

About the Authors

Anastasia Lada is a doctoral researcher at the Vrije Universiteit Brussel (VUB), specializing in idiom comprehension in healthy ageing and Alzheimer's disease. Her research focuses on figurative language processing in both ageing populations and individuals with neurological conditions, including Alzheimer's disease and acquired aphasia.

Ifigeneia Dosi is an Assistant Professor of Applied Linguistics at Democritus University of Thrace, who works on bilingual language development. She was awarded an IKY scholarship for her research on word definition strategies in children with language disorders. With a focus on bilingual children's language and cognitive development, she has published in high-impact scientific journals. Additionally, she has provided training for educators working with refugees and heritage speakers. Her research interests center on studying the language development of bilingual children, examining the interplay of cognitive and environmental factors in both typical and atypical development.

Philippe Paquier is a full professor emeritus of psycho- and neurolinguistics at VUB, and an associate professor emeritus of neurolinguistics and aphasiology at ULB. His main research interests focus on aphasia and other neurological speech and language disorders in children (incl. Landau-Kleffner Syndrome and Cerebellar Mutism Syndrome).

Christina Manouilidou is an Associate Professor of Psycholinguistics and Neurolinguistics. Her research interests include (1) cross-linguistic investigations of lexical access and representation of complex words (2) cognitive and linguistic impairments in aging populations with chronic neurodegenerative diseases, and (3) language intervention and therapy in developmental and acquired language disorders by employing behavioral and neuromodulation techniques.

Stefanie Keulen is an associate professor of psycholinguistics and neurolinguistics at VUB and a lecturer of aphasiology at ULB in Belgium. Her research interests are (primary progressive) aphasia, motor speech disorders, and neuromodulation.

Appendix A

A list of 199 idioms rated in familiarity and predictability. The list below provides the mean score and standard deviation for each idiomatic expression. For familiarity the means range from 0 to 5 while for predictability the means range from 0 to 1.

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
τυφλός στην αγάπη (= είναι πολύ ερωτευμένος)	4.46	1.062	.49	.504
παίρνω τον αέρα (= αποκτώ θάρρος)	4.62	1.011	.80	.408
κάνω τ'αλατιού (=δέρνω πολύ)	2.54	2.085	.42	.496
περί ανέμων και υδάτων (=μιλάω γενικά χωρίς ουσία)	4.68	.812	.98	.126
άρον άρον (= πολύ βιαστικά)	4.88	.650	.98	.126
τεντώνω τα αυτιά μου (=ακούω προσεκτικά)	4.55	.902	.98	.126
φέρω βαρέως (=νιώθω αμήχανα εξαιτίας μείωσης της αξιοπρέπειας μου)	2.03	1.887	.18	.368
ανοίγω το βήμα μου (=περπατάω γρήγορα/ επιταχύνω)	4.34	1.176	.72	.455
φέρνω βόλτα (=σαγηνεύω/πείθω)	4.72	.696	.11	.317
την βρίσκω (= διασκεδάζω/περνάω καλά)	4.69	.748	.71	.463
βγαίνω γελασμένος (= απατούμαι)	4.34	.889	.00	.000
με το γλυκό (=προσεκτικά και απαλά)	2.89	1.838	.00	.000
τραβάω γραμμή (=πηγαίνω κατευθείαν)	4.15	1.278	.05	.215
χύνω κροκοδείλια δάκρυα (= προσποιούμαι συγκίνηση)	4.43	1.212	.94	.246
βρίσκω το δίκιο μου (= δικαιώνομαι)	4.80	.833	.97	.126
παίρνω τον ολισθηρό δρόμο (=παραστρατώ)	3.80	1.471	.86	.353
πάω από εκεί που ήρθα (= διώχνομαι)	4.62	.979	.98	.126
σκάω την ζαχαρένια μου (=ανησυχώ/στενοχωριέμαι)	4.45	1.046	.20	.396
δεν με βλέπει ο ήλιος (= δεν βγαίνω από το σπίτι)	4.68	.731	.94	.215
πέφτω του θανατά (=το παίρνω κατάκαρδα)	4.43	1.299	.78	.419
φέρνω σε δύσκολη θέση (=προκαλώ αμηχανία)	4.92	.322	1.00	.000

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
δεν έχω ιδέα (=δε γνωρίζω)	4.89	.472	.86	.336
βρίσκομαι επί τα ίχνη (=είμαι στον σωστό δρόμο να βρω κάτι)	2.98	1.883	.05	.215
αφήνω κάγκελο (= εκπλήσσω)	4.66	.776	.05	.215
σφίγγει ο καιρός (=καλοκαιριάζει)	4.06	1.391	.43	.499
στα καλά καθούμενα (= ξαφνικά)	4.88	.650	.97	.177
βάζω κάλπη (=θέτω υποψηφιότητα)	2.17	1.710	.02	.126
με μισή καρδιά (= διστακτικά)	4.83	.575	1.00	.000
δεν μου καίγεται καρφί (= αδιαφορώ)	4.91	.384	.98	.126
παίρνω κάτι κατάκαρδα (=στεναχωριέμαι πολύ για κάτι)	4.78	.800	.78	.408
χάρις στην κεραμιδιά μου (=εξαιτίας του έρωτα μου)	.97	1.479	.00	.000
ότι μου κατέβει στο κεφάλι (=οτι σκεφτώ)	4.85	.565	.85	.368
κάνω κάποιον κέφι (=συμπαθώ κάποιον / τον εγκρίνω)	4.23	1.183	.51	.504
στέκομαι κλαρίνο (= είμαι σε στάση προσοχής)	3.48	1.760	.05	.215
σπάει κόκαλα (=είναι ανυπόφορο)	3.82	1.457	.22	.396
ένας κόμπος έχει σταθεί στον λαιμό μου (= δυσκολεύομαι να μιλήσω)	4.72	.625	.94	.246
κάνω κόρτε (=φλερτάρω)	1.65	1.940	.02	.126
δε σηκώνω κουβέντα (= είμαι αμετάπειστος)	4.89	.359	.63	.485
από κούνια (=ανέκαθεν)	4.58	.950	.29	.463
γίνομαι κουφός και μουγκός (= προσποιούμαι οτι δεν ακούω κι δεν αποκρίνομαι)	3.52	1.572	.12	.336
η τύχη μου κρέμεται από κάτι (=εξαρτώμαι από κάτι)	4.62	1.041	.37	.485
την βγάζω λάδι (=αθωώνομαι, δεν τιμωρούμαι)	4.52	1.077	.35	.485
δεν βγάζω λέξη (=δεν μιλάω)	4.75	.685	.03	.177
γίνομαι λιάδα (ή λιάρδα) (=μεθάω)	3.91	1.748	.05	.177
μασημένα λόγια (=υπεκφυγές, περιστροφές στα λόγια)	4.22	1.244	.02	.126

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
έρχομαι στα λόγια (=λογομαχώ)	4.12	1.420	.03	.177
αφήνω στα κρύα του λουτρού (=παρατάω)	4.66	.957	.95	.215
με πιάνει λύσσα (= νιώθω έντονη σεξουαλική επιθυμία)	4.43	1.015	.05	.215
λέω το μακρύ μου και το κοντό μου (=μιλάω αυθαίρετα)	4.74	.815	.98	.126
τα κάνω όλα μαντάρα (=δημιουργώ αναταραχή)	4.74	.815	.28	.447
βλέπω με μισό μάτι (=δεν εμπιστεύομαι)	4.80	.642	.97	.177
ανοίγω τα μάτια κάποιου (=βοηθάω κάποιον να συνειδητοποιήσει την πραγματικότητα)	4.86	.464	.91	.296
έχω τα μάτια μου τέσσερα (=προσέχω πολύ)	4.26	1.266	.00	.000
κάνω τα στραβά μάτια (= παραβλέπω)	4.83	.486	.98	.126
μου κόβεται το γέλιο μαχαίρι (= σταματάω απότομα να γελώ)	4.80	.506	.51	.504
είμαι μέσα στα πράγματα (=γνωρίζω πολύ καλά)	4.85	.537	.58	.499
μαζί μιλάμε και χώρια καταλαβαινόμαστε (= δεν συνεννοούμαστε)	4.63	.876	.69	.469
πουλάω μούρη (=υπερηφανεύομαι)	4.68	.868	.63	.485
τρώω τα μούτρα μου (=αποτυγχάνω)	4.85	.507	.85	.353
με παίρνει η μπάλα (=βρίσκω τον μπελά μου)	4.85	.441	.95	.215
τρώω το μπερντάχι μου (=με δέρνουν)	1.15	1.813	.06	.246
κολλάω μια μπουνιά (= γρονθοκοπώ)	3.77	1.656	.48	.504
μου γυρίζει τα μυαλά (=αλλάζω γνώμη για αυτή)	4.46	1.147	.75	.441
τινάζω τα μυαλά μου στον αέρα (= αυτοκτονώ με όπλο)	4.82	.583	.98	.126
κόβει το μυαλό μου (=είμαι έξυπνος)	4.88	.545	.78	.408
κάτι στριφογυρίζει στο μυαλό μου (=σκέφτομαι κάτι)	4.74	.644	.54	.502
η ιστορία μύριζε (=ήταν ύποπτη)	3.74	1.623	.05	.215

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
δεν βλέπω πέρα από τη μύτη μου (=δεν είμαι διορατικός)	4.71	.744	.80	.408
μου πέφτει η μύτη μου και δεν σκύβω να την πάρω (=είμαι υπερήφανος)	3.18	1.845	.08	.272
με τσιμπάει μύγα (=αλλάζω ξαφνικά συμπεριφορά)	4.77	.766	.97	.177
πηγαίνω με τα νερά κάποιου (= δέχομαι τις απόψεις κάποιου για να μην εξαγριωθεί)	4.91	.423	.92	.272
κάνω το νερό μου (= ουρώ)	1.86	1.911	.03	.177
κάνω πόλεμο νεύρων σε κάποιον (= εκνευρίζω κάποιον)	4.18	1.379	.51	.504
μου βάζουν νέφτι (=βιάζομαι υπερβολικά)	3.31	1.870	.29	.455
βάζω με το νου μου (= σκέφτομαι)	4.72	.839	.45	.501
κοιτάω το ταβάνι (=είμαι αδρανής)	4.55	1.146	.69	.469
πηγαίνω ντουγρού (=πηγαίνω κατευθείαν)	4.65	.891	.11	.296
έχω το κοκαλάκι της νυχτερίδας (=είμαι τυχερός)	3.88	1.816	.77	.429
κάποιος με έχει ξεγραμμένο (=κάποιος θεωρεί οτι δεν έχω ελπίδες)	4.78	.573	.38	.490
ξερνάω χολή (=μιλάω με κακία)	4.29	1.296	.15	.368
μελανιάζω κάποιον στο ξύλο (= χτυπάω κάποιον πολύ)	4.63	.928	.75	.429
όνειρο θερινής νυκτός (=απραγματοποίητο όνειρο)	3.88	1.772	.78	.419
έχω κάποιον στα όπα όπα (= φροντίζω κάποιον ιδιαίτερα)	4.86	.583	.62	.490
πέφτω από τον ουρανό (=μένω έκπληκτη)	4.72	.927	.72	.447
το τραβάει ο οργανισμός κάποιου (=το αντέχει κάποιος)	4.62	.979	.89	.317
σπάω τον πάγο (=ξεπερνάω την αρχική αμηχανία)	4.78	.696	.97	.177
πάει να πεί (=δηλαδή)	4.58	.998	.92	.272
χάνω το παιχνίδι (=αποτυγχάνω)	4.78	.625	.20	.396
παίρνω κάτι πίσω (=ανακαλώ)	4.77	.702	.91	.296
είναι για τα πανηγύρια (=είναι γελοίος)	4.77	.606	.32	.469
έρχονται τα πάνω κάτω (=έχει προκληθεί αναστάτωση)	4.88	.415	.98	.126
γίνομαι βαπόρι (=εξοργίζομαι)	2.72	1.807	.00	.000

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
τρώω το παραμύθι (=πιστεύω σε κάτι)	4.83	.486	.35	.485
ανοίγω παρτίδες (= έχω σχέσεις)	4.51	.986	.46	.503
κάνω πάσα (=δίνω)	4.20	1.427	.35	.481
πατάω στη γη (=είμαι ρεαλιστής)	4.62	.860	.86	.353
κάνω πεζοδρόμιο (=να εκπορνεύομαι)	3.38	2.044	.25	.429
βάζω στον πειρασμό (=δελεάζω)	4.65	.874	.40	.496
είμαι πέρδίκι (=είμαι καλά)	4.86	.556	.69	.469
πάνε όλα περίπατο (=χάνονται)	4.31	1.117	.08	.272
τινάζω τα πέταλα (=πεθαίνω)	4.03	1.610	.66	.481
μένω πετρωμένος (=μένω ακίνητος)	3.54	1.640	.00	.000
που πέφτει (=που βρίσκεται)	4.05	1.556	.43	.496
μου πηγαίνει (=μου ταιριάζει)	4.75	.952	.94	.246
με πιάνει κάτι (=με αγγίζει κάτι)	4.58	.827	.11	.317
είμαι πουλί της πιάτσας (=είμαι έμπειρος στις συναλλαγές)	3.22	1.892	.15	.368
έχω τυφλή πίστη (=έχω απόλυτη εμπιστοσύνη)	4.54	.937	.11	.317
με έκοψε πλάγια (=με παρατήρησε με το πλάι του ματιού)	2.09	1.826	.03	.177
κάποιος μου κάνει χοντρή πλάκα (=κάποιος με ξεγελάει)	4.63	.928	.22	.419
έχω κάποιον στην πλάτη μου (=κάποιος με επιβαρύνει)	4.26	1.163	.74	.439
ξύνω παλιές πληγές (=θυμίζω παλιές πληγές)	4.83	.698	.89	.317
βάζω πλώρη (=κατευθύνομαι)	4.20	1.252	.38	.485
κάνω την ζωή κάποιου ποδήλατο (= ταλαιπωρώ κάποιον)	4.29	1.271	.05	.126
πατάω το πόδι μου (=πηγαίνω)	4.17	1.398	1.00	.000
μπερδεύομαι στα πόδια κάποιου (=αποσπώ κάποιον από την εργασία του)	4.78	.673	.92	.272
που σε πονεί και που σε σφάζει (=κάποιον τον χτυπούν πολύ)	3.00	2.000	.34	.475
κατεβάζω μερικά ποτηράκια παραπάνω (=μεθάω)	4.75	.685	.74	.447

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
μέσα στα πράγματα (=είναι γνώστης)	4.65	.975	.62	.493
καθώς πρέπει (=αξιοπρεπής)	4.85	.537	.98	.126
προσγειώθηκε ανώμαλα (=ξαφνιάστηκε αρνητικά)	4.12	1.244	.03	.126
πιάνει πουλιά στον αέρα (=έχει ταχεία αντίληψη)	4.83	.575	.98	.126
γίνομαι πυρ και μανία (=εξοργίζομαι)	4.69	.828	.94	.246
μένω ρέστος (=ξεμένω)	2.98	2.080	.03	.126
το ρίχνω στο ξεκάρφωτο (=λέω κάτι όταν δεν το περιμένει κάποιος)	3.91	1.355	.03	.177
βγήκαν ρολόι (=ήταν στην ώρα τους)	3.68	1.511	.03	.177
βγαίνω από τα ρούχα μου (=εξοργίζομαι)	4.88	.673	.92	.246
μου τρέχουν τα σάλια (=επιθυμώ κάτι που δεν μπορώ να αποκτήσω)	4.97	.248	.98	.126
παίρνω σβάρνα (=παρασύρω)	4.75	.830	.48	.504
το κλίμα σηκώνει (=οι συνθήκες είναι κατάλληλες)	4.65	.799	.48	.503
πέφτει μαύρη σιωπή (=επικρατεί απόλυτη σιωπή)	3.88	1.409	.26	.447
τα σκατώνω (=αποτυγχάνω)	4.62	.896	.12	.336
κάνω σε κάποιον σκηνή (=τσακώνομαι για λόγους ζηλοτοτυπίας)	4.91	.341	.66	.481
τρώει κάποιον το μαύρο σκοτάδι (=πεθαίνω)	3.63	1.567	.12	.336
σέρνω σε κάποιον όσα δε μαζεύει η σκούπα (=λέω σε κάποιον πολλά άσχημα πράγματα)	2.31	1.819	.22	.419
δεν είναι σόι (=δεν είναι καλής ποιότητας)	4.22	1.340	.17	.383
κλείνω κάποιου το σπίτι (=χάνω την περιουσία κάποιου)	4.49	1.033	.88	.336
κατεβάζω σταυρούς και Παναγίες (=βρίζω τα Θεία)	4.72	.839	.51	.504
γίνομαι στήλη (=παγώνω)	2.60	1.910	.06	.215
βάζω κάτι στο στόμα μου (=τρώω λίγο)	4.63	1.009	.83	.383
στον λαιμό μου κάθεται (=αντιπαθώ κάποιον)	4.85	.537	.98	.126

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
πηγαίνει στράφι (=χάνεται)	4.88	.415	.60	.493
με φέρνουν στα συγκαλά μου (=με κάνουν να συνέλθω)	4.66	.735	.55	.499
μου έχει πάρει τα συλλοϊκά μου (=με έχει καταγοητεύσει)	1.72	1.883	.00	.000
τρώω σφαλιάρα (=με χαστουκίζουν)	4.66	.735	.11	.296
το έχω μέσα στο σώμα μου (=είναι στη φύση μου)	3.42	1.638	.03	.177
γίνομαι σε κάποιον ταγάρι (=ενοχλώ κάποιον)	1.29	1.465	.00	.000
μου έρχεται ταμπλάς (=μένω έκπληκτος)	4.22	1.463	.17	.383
χορεύω κάποιον στο ταψί (=ταλαιπωρώ κάποιον)	4.71	.843	.97	.177
τρέχω με τα τέσσερα (=υπακούω)	2.37	1.884	.03	.177
το και το (=με λεπτομέρειες)	4.52	1.200	.83	.383
εκτός τόπου και χρόνου (=εκτός πραγματικότητας)	4.88	.451	1.00	.000
τράβα με και ας κλαίω (=κάνω κάτι παρόλο που μου είναι δυσάρεστο)	4.77	.679	.98	.126
με πιάνει μια τρέλα (=λειτουργώ παράλογα)	4.69	.846	.72	.455
θα τον φάμε (=θα επικρατήσουμε, θα τον νικήσουμε)	4.14	1.102	.05	.215
τα λέω στην τρίχα (=τα διηγούμαι λεπτομερώς)	2.62	2.074	.08	.272
θα γυρίσει ο τροχός (=τα πράγματα θα αλλάξουν)	4.80	.617	.88	.336
μου έφαγε (=μου έκλεψε)	3.62	1.578	.23	.408
τσάμπα και βερεσέ (=χωρίς λόγο)	4.12	1.463	.57	.501
είναι τσιμπημένος (=είναι ερωτευμένος)	4.60	.862	.14	.353
βράζει το τσουκάλι (=εξασφαλίζω τα προς το ζην)	2.25	1.888	.05	.215
μιλάω με την τύχη μου (=είμαι τυχερή)	3.82	1.550	.65	.481
σκάει από υγεία (=είναι απόλυτα υγιές)	4.66	.834	.78	.419
με παίρνει ο ύπνος (=αποκοιμιέμαι)	4.80	.592	1.00	.000
παίρνω τα ύψη (=ανεβαίνω πολύ ψηλά)	1.98	1.781	.05	.215

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
με πιάνει η φαγούρα (=επιθυμώ)	4.03	1.323	.03	.177
κρατώ σε κάποιους το φανάρι (=διευκολύνω την ερωτική συνεύρεση)	5.00	.000	1.00	.000
βρίσκομαι φάτσα (=είμαι απέναντι)	3.05	1.789	.03	.177
την φέρνω σε κάποιον (=εξαπατώ κάποιον)	4.63	.858	.75	.429
είμαι φέσι (= είμαι μεθυσμένος)	3.60	1.818	.00	.000
βγάζω το φίδι από την τρύπα (=βρίσκω την λύση)	4.74	.776	.94	.246
παίρνω κάτι φιλοσοφικά (=αντιμετωπίζω κάτι με ηρεμία)	3.02	1.644	.00	.000
βγάζω στη φόρα (=κοινοποιώ)	4.88	.451	.95	.215
στην φούρια που με πιάνει (=στην βιασύνη μου)	3.38	1.588	.12	.336
έχω σώας τας φρένας (=είμαι λογική)	4.85	.507	.97	.177
τα φτιάχνω με (=έχω δεσμό)	4.45	1.347	.18	.383
δεν κάνω ούτε για φτύσιμο (=δεν αξίζω)	4.72	.650	.06	.246
τα φυσάει (=είναι πλούσιος)	3.60	1.703	.26	.439
μπήγω μια φωνή (=φωνάζω)	4.23	1.308	.74	.439
ανάβω φωτιά στην καρδιά κάποιου (= σαγηνεύω κάποιον)	4.51	.904	.72	.447
γίνεται χάβρα των Ιουδαίων (=γίνεται πολλή φασαρία)	2.52	2.251	.48	.502
τα χαλάσαμε (=διαφωνήσαμε)	4.72	.820	.77	.432
ο κόσμος να χαλάσει (=ό,τι και εάν συμβεί)	4.88	.484	.92	.272
φτάνω χαμηλά (=ξεπέφτω)	3.91	1.588	.18	.383
τα 'χασα (=ξαφνιάστηκα, σάστισα)	4.83	.486	.72	.455
βαστάω χαρακτήρα (=μένω πιστός)	4.02	1.305	.02	.126
δείχνω τα χαρτιά μου (=αποκαλύπτω τις προθέσεις μου)	4.48	1.062	.37	.490
είμαι όλο χαχαχά και χουχουχού (=είμαι χαζοχαρούμενος)	4.86	.496	.89	.296
απλώνω χέρι (=κλέβω)	4.42	1.158	.94	.246
βάζω το χέρι μου στη φωτιά (=είμαι απόλυτα σίγουρος)	4.95	.276	.98	.126

IDIOMS	Familiarity	Std. Deviation	Predictability	Std. Deviation
οτι περνάει από το χέρι μου (=οτι μπορούσα/είχα την δυνατότητα)	4.78	.625	.98	.126
έρχομαι στα χέρια (=διαπληκτίζομαι)	4.60	1.028	.48	.502
κολυμπάω στο χρυσάφι (=είμαι πάμπλουτος)	4.23	1.284	.02	.126
ρίχνω σε κάποιον χυλόπιτα (=απορρίπτω κάποιον)	4.92	.444	.29	.463
τι ψάρια πιάνω (=τι αντιλαμβάνομαι/καταλαβαίνω)	4.46	1.147	.94	.246
τα ψέλνω από την καλή (=επιπλήττω)	4.48	1.133	.48	.502
ψήνω κάποιον (=πείθω κάποιον)	4.71	.701	.25	.439
για ένα κομμάτι ψωμί (=πολύ φθηνά)	4.62	.979	.95	.215
για ψύλλου πήδημα (=με το παραμικρό)	4.63	.993	.89	.317
της κακιάς ώρας (=κακής ποιότητας)	4.88	.484	.97	.177