

On the relevance of the syntactic flexibility of an idiom for its recognition: Experimental evidence from Polish

Hanna Kędzierska, Wojciech Witkowski, Joanna Błaszczak, Dorota Klimek-Jankowska, Piotr
Gulgowski (University of Wrocław)

Abstract The article is a contribution to a long-standing discussion on how idioms are represented and accessed in the mental lexicon. More specifically, in a timed cloze response study we investigate high and low syntactic flexibility idioms in Polish in order to find out whether the degree of syntactic flexibility influences the ease and time of idioms' recognition. By doing so we contribute to the question of whether idioms are differently represented in the lexicon depending on their syntactic flexibility, as suggested by Nunberg et al. (1994) and Gibbs and Nayak (1989), or whether all idioms independent of their syntactic flexibility are represented lexically in the same hybrid way, as suggested by Cutting and Bock (1997) and Sprenger et al. (2006). The results of our study support the latter view.

Keywords: idiom recognition, syntactic flexibility, hybrid lexical representation of idioms, superlemma

1 Introduction

In this paper we contribute new experimental facts from Polish shedding light on the discussion related to the questions of (i) how idioms are represented and accessed in the mental lexicon and (ii) whether this depends on the degree of their syntactic flexibility. In early generative approaches, all idioms were viewed as non-compositional units whose meanings are not computed based on the meanings and syntactic relations of their components (see Katz and Postal 1963; Katz 1973; Chomsky 1980; Machonis 1985; Glasbey 2003, among others). Similarly, in early psycholinguistic studies all idioms were assumed to be stored and retrieved as elementary lexical units ("long words") from the mental lexicon (see Bobrow and Bell 1973; Swinney and Cutler 1979, among others). This unitary view was later challenged by theoretical linguists and psycholinguists (Nunberg 1979; Wasow et al. 1983; Gibbs et al. 1989a, b; Cacciari and Glucksberg 1991; Müller 2000; Cieślicka 2012; Leivada 2017, among others), who pointed out that idioms do not constitute a uniform class and they can be divided into decomposable (e.g., *to lay down the law*) and non-decomposable ones (e.g., *to chew the fat*). The property of decomposability was associated with the semantic relatedness between the idiom's figurative and literal meanings (Gibbs et al. 1989a, b; Cacciari and Glucksberg 1991; see also Geeraerts 1995). This division was later correlated with syntactic flexibility by Nunberg et al. (1994) (see also Gibbs and Nayak 1989). More specifically, they suggested that the more semantically decomposable an idiom is the more syntactically flexible it is expected to be. Furthermore, Nunberg et al. (1994) postulate a division of idioms into (i) idiomatically combining expressions such as, for example, *spill the beans*, whose meanings are

identifiable from the meanings of their parts (for example, when we interpret the idiom *spill the beans*, *spill* is associated with the meaning of *to divulge* and *the beans* is semantically associated with *the information that is divulged*) and (ii) idiomatic phrases such as, for example, *kick the bucket*, *saw logs* whose meanings are not retrievable from the meanings of their components (for instance, when we interpret the idiom *kick the bucket* as *to die*, the concept of ‘dying’ cannot be attributed to the meanings of the idiom’s constituents). Nunberg et al. (1994) argue that only idiomatic phrases, being non-decomposable and syntactically frozen, are stored in the lexicon as complete phrases, which would suggest that the literal meaning of their constituents cannot be accessed and given that such idioms have no internal structure, they cannot be syntactically modified. By contrast, in the case of idiomatically combining expressions, the literal meaning of their constituents is accessible and as such they can be syntactically modified, e.g., by means of adjectives or relative clauses, or they can be omitted in elliptical constructions (VP ellipsis), or can be used as antecedents for pronouns, as shown in (1)-(5) respectively (see Nunberg et al. 1994: 500-502).

- (1) *leave no legal stone unturned*
- (2) *Your remark touched a nerve that I did not even know existed.*
- (3) *Those strings, he wouldn’t pull for you.*
- (4) *My goose is cooked but yours isn’t.*
- (5) *We thought that tabs were being kept on us but they weren’t.*

This means that only idiomatically combining expressions are syntactically flexible and, as suggested by Nunberg et al. (1994), only in their case the literal meaning of the single lemmas can be accessed. The latter claim is not compatible with the results of recent psycholinguistic experiments showing that literal word meanings of all idioms (irrespective of their syntactic flexibility and decomposability) become active during idiom comprehension (Cacciari and Tabossi 1988) and production (Cutting and Bock 1997; Sprenger et al. 2006), which suggests that an idiom’s lexical entry is linked to its constituent literal meanings. More precisely, Cutting and Bock (1997) in the experimental study on idiom blends tested the hypothesis of Gibbs and Nayak (1989) that the lexical representation of semantically decomposable idioms is less rigidly specified and more susceptible to change than that of non-decomposable idioms. If this hypothesis is correct, the prediction would be that semantically decomposable idioms should result in more idiom blends in the error elicitation task. Contrary to this prediction, the error rates in their experiment were not correlated with the degree of idiom’s decomposability. Based on this evidence, Cutting and Bock (1997) argue that the lexical representations of decomposable and non-decomposable idioms are the

same when they enter into the speech production process. In another experiment, Peterson et al. (2001) reported a syntactic priming effect for idiomatic phrases independent of the degree of structural flexibility of a given idiom (see Sprenger et al. 2006). More precisely, the participants saw sentences ending with the beginnings of idiomatic phrases but the preceding context supported either their idiomatic or literal continuation. At the end of each incomplete idiomatic phrase, the participants saw an unrelated noun or verb that they were asked to read aloud as quickly as possible. Peterson et al. (2001) observed that the idiomatic contexts set up an expectation for the noun category and this effect was independent of their syntactic flexibility. This suggests that we have access to the information about the syntactic category of the idioms' constituents and it is not influenced by the degree of the idiom's syntactic frozenness and/or its decomposability. Based on these findings, Cutting and Bock (1997) and Sprenger et al. (2006) suggested that all idioms have a hybrid lexical representation in that at one level, idioms have unitary entries corresponding to their non-literal meanings but at another level, they are represented by simple lemmas of their single constituents. In both Cutting and Bock's (1997) and Sprenger et al.'s (2006) accounts it is assumed that syntactic information about a given idiomatic phrase is lexically encoded. Cutting and Bock (1997) postulate that the idiom's lexical representation is linked to its pre-fabricated frame understood as a phrase structure with open slots that can be filled with simple lemmas that are activated by the idiom's lexical concept node (see Figure 1). In Sprenger et al.'s (2006) approach it is assumed that the mental representation of an idiom (the so called superlemma, see Figure 2) also contains information about its syntactic idiosyncrasies (e.g., no passive form, no adjectival modification, see Figure 3).

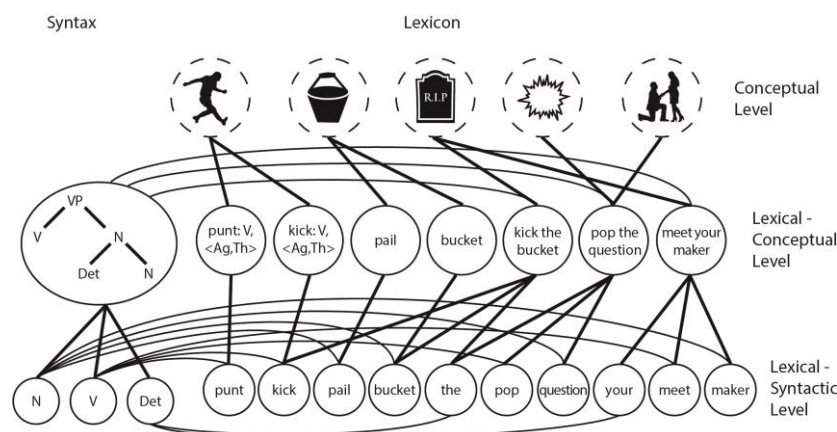


Figure 1. Idiomatic representation in the mental lexicon (Cutting and Bock 1997: 67)

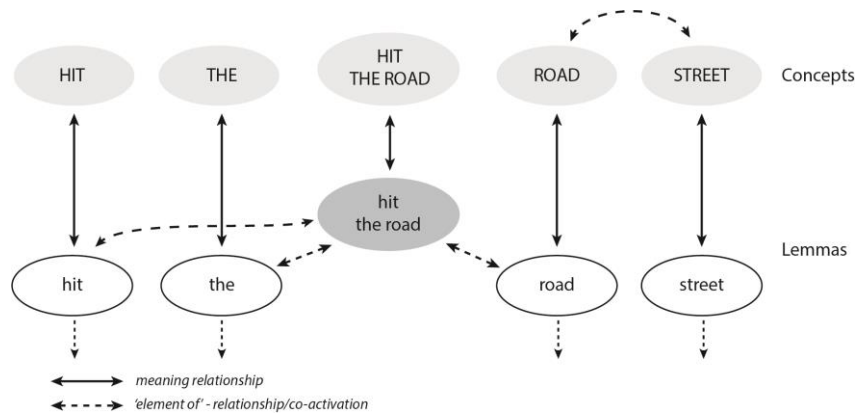


Figure 2. Idiomatic representation of an idiom *to hit the road* (Sprenger et al. 2006: 176)

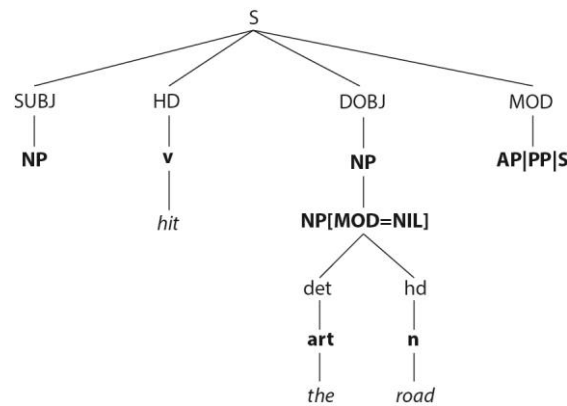


Figure 3. Information represented in the superlemma for *hit the road*. The lemmas *hit*, *the*, and *road* are obligatory components. Modification of *road* (e.g., **hit the icy road*) is not allowed (see Sprenger et al. 2006: 177)

Following Cacciari and Tabossi (1988), Sprenger et al. (2006) assume that the unitary representation of an idiom is activated once the comprehender has processed some information allowing for a unique identification of a given idiom, the so-called idiom's key or idiom recognition point (see Vespignani et al. 2010). Before the idiom recognition point is reached, the comprehenders process the idiom's components literally. The concept of the idiom recognition point is intuitively very convincing, however still little is known about which factors may influence it. In her recent experimental study, Kędzierska (2018) showed that the idiom's recognition point is reached faster in a facilitating context. In the present study, we ask a related question and investigate the influence of the syntactic flexibility of an idiom on its recognition point, which in the light of the discussed studies has a potential to contribute to the question of whether idioms are differently represented in the lexicon depending on their syntactic flexibility, as suggested by Nunberg et al. (1994) and Gibbs and Nayak (1989), or whether all idioms independent of their syntactic flexibility are represented lexically in the same hybrid way, as suggested by Cutting and Bock (1997) and Sprenger et al. (2006). Based on Nunberg et al.'s (1994) claim that only syntactically frozen idioms are stored as chunks and syntactically flexible idioms are composed, one would predict that the former should be

recognized and retrieved faster. In contrast, based on Sprenger et al.'s (2006) account assuming that the information about the idiom's syntactic properties (their being more or less syntactically flexible) is accessed after the comprehender reaches the idiom's recognition point leading to the activation of the idiom's lexical representation, the syntactic flexibility should not affect the time needed to recognize an idiom. With these competing predictions in mind, we designed a timed cloze response study in which we tested whether the idiom's recognition point depends on the degree of its syntactic flexibility.

2 The present study

2.1 Material

For the purpose of the study a list of 33 Polish idioms was created. All the selected idioms had a similar syntactic structure, that is, all the idioms were verb phrases headed by transitive verbs followed by a complement and a prepositional phrase (V+DP_{OBJ}+P+DP_{PREP_OBJ}¹), as illustrated in (6). Moreover, all the selected idioms had both plausible non-idiomatic (7a) and idiomatic meanings (7b).

- (6) *mieć asa w rękawie*
 have.INF ace.SG.ACC in sleeve.SG.LOC²
 'to have an ace up your sleeve'

- (7) a. *Gdy graliśmy w karty, cały czas miałem asa w rękawie.*
 when play.PST.IPFV.1PL in card.PL.ACC [whole time].SG.ACC
 have.PST.1SG.M ace.SG.ACC in sleeve.SG.LOC
 'While we played cards, I had an ace up my sleeve all the time.'
- b. *Byliśmy na skraju bankructwa, ale miałem jeszcze asa w rękawie.*
 be.PST.1PL on verge.SG.LOC bankruptcy.SG.GEN but have.PST.1SG.M
 still ace.SG.ACC in sleeve.SG.LOC
 'We were about to go bankrupt but I still had an ace up my sleeve.'

¹ We use the general term "Determiner Phrase" (DP) for ease of reference without commitment to the precise structure of nominal phrases, which is especially problematic in article-less languages such as Polish. "DP_{PREP_OBJ}" stands for a DP which serves as a complement to a preposition (e.g., *my sleeve* in the PP *up my sleeve* in (7)).

² The abbreviations used in the glosses are based on the Leipzig Glossing Rules (see <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>).

In order to estimate the syntactic flexibility of the selected 33 idioms we chose 13 syntactic flexibility tests (some of which were earlier reported in the related studies; see Nunberg et al. 1994; Stone 2016; Wierzba 2016).

2.2 Syntactic flexibility tests – a corpus study

The syntactic flexibility tests were conducted on the basis of the data available in Araneum Polonicum Maius (Polish, 15.02) (Benko 2014). First, it was checked whether a given idiom could be found in the corpus in the modified form under question (e.g., negative form, passivized form, modal form, etc.) without losing its figurative meaning. Examples from outside the corpus were provided only if no modified version of an idiom had been found and the independent judgements of three Polish native speakers demonstrated that such a modification is possible. The 13 tests, which have been administered in order to determine the syntactic flexibility of 33 idioms used in the study, are summarized below.

Test 1: NEGATION

The aim of the test was to determine whether a negative form of an idiom is possible to construct.

- (8) *Lepiej poczekać kilka tygodni i nie dolewać oliwy do ognia.*
 better wait.INF [several week].PL.GEN and NEG add.INF olive.SG.GEN to
 fire.SG.GEN
 ‘It is better to wait several weeks than to escalate the problem.’

Test 2: PASSIVIZATION

The test was conducted in order to verify whether a given idiom may be passivized.

- (9) *Dzisiaj o godzinie 14 wszystkie karty zostaną wyłożone na stół.*
 today at hour.SG.LOC 14 all card.PL.NOM get.PRS.PFV.3PL
 lay.PTCP.PL on table.SG.ACC
 ‘Today at 2 p.m. we will finally be honest about our intentions.’

Test 3: MODAL VERBS

The goal of the test was to establish whether a particular idiom can be preceded by a modal or modalized verb. A set of modal verbs used in the queries included: *musieć* (‘must’), *mieć* (‘have to’), *móc* (‘can’), *potrafić* (‘can/be able to’), *chcieć* (‘want’), *woleć* (‘prefer’).

- (10) *Kandydat na prezydenta nie może chować*
 candidate.SG.M.NOM on president.SG.GEN NEG can.3SG hide.INF
głowy w piasek.
 head.SG.GEN in sand.SG.ACC
 ‘A presidential candidate cannot avoid responsibility.’
- (11) *Z tego powodu wolę trzymać rękę*
 with this.SG.GEN reason.SG.GEN prefer.PRS.1SG keep.INF hand.SG.ACC
na pulsie.
 on pulse.SG.LOC
 ‘For this reason, I prefer to stay alert.’

Test 4: VP-EXTERNAL (ADVERBIAL) MODIFICATION

The test was conducted in order to check whether a given idiom may be modified by a higher adverbial (see Cinque 1999); e.g., a temporal or frequentative adverbial.

- (12) (...) *czy czytałaś wywiad z M. Gretkowską,*
 Q read.PST.IPFV.2SG.F interview.SG.ACC with M. Gretkowska.SG.F.INS
która wsadziła ostatnio kij w mrowisko³?
 who.SG.F.NOM put.PST.PFV.3SG.F recently stick.SG.ACC in anthill.SG.ACC
 ‘(...) Have you read an interview with M. Gretkowska who has recently made a lot of people angry?’

Test 5: VP-INTERNAL MODIFICATION

The test was conducted in order to check whether a given idiom may be internally modified (or, in other words, whether a manner adverb can be merged with VP *or* whether an adjective can be merged with DP).

- (13) *Zamiast nam pomóc, inteligenci budują*
 instead 1PL.DAT help.INF intellectual.PL.NOM build.PRS.IPFV.3PL
wyrafinowane zamki na piasku⁴.
 sophisticated.PL.ACC castle.PL.ACC on sand.SG.LOC
 ‘Instead of helping us, the intellectuals dream about things which are impossible to realize.’

³ *Wsadzić kij w mrowisko*: literally, ‘to put a stick into the ant hill’. The figurative meaning of this idiom involves making somebody angry by bringing up some controversial or unpleasant issue.

⁴ *Budować zamki na piasku*: literally, ‘to build castles on the sand’. The figurative meaning of this idiom involves dreaming about unlikely events.

- (14) *Autorce* *trzeba było* *łopatologicznie* *wyłożyć*
 author.SG.F.DAT need be.PST.3SG.N clearly (colloquial) lay.INF
*kawę na ławę*⁵.
 coffee.SG.ACC on table.SG.ACC
 ‘As for the author, we had to very clearly tell her what to do.’

Test 6: VERB-ELLIPSIS

The aim of the test was to verify whether the verbal head of an idiom may be elided.

- (15) *Czy jesteśmy skazani na hasła i zamki na*
 Q be.PRS.1PL condemned on password.PL.ACC and castle.PL.ACC on
piasku?
 sand.SG.LOC
 ‘Are we condemned to empty words and visions which are very unlikely to come true?’

Test 7: VERB SUBSTITUTION

The purpose of the test was to determine whether the verbal head of an idiom may be substituted with another verb. For example, the verb *dolewać* (‘pour’) in *dolewać oliwy do ognia* (lit. ‘to pour fuel to the fire’) may be substituted with such verbs as *dłożyć* (‘to pile on’) or *dodawać* (‘to add’).

- (16) a. *W tym momencie rząd dolewa*
 in this.SG.LOC moment.SG.LOC government.SG.NOM
 pour.PRS.IPFV.3SG
oliwy do ognia.
 olive.SG.GEN to fire.SG.GEN
 ‘At the moment, the government is escalating the problem.’
- b. *Oliwy do ognia dłożyło, że firma*
 olive.SG.GEN to fire.SG.GEN pile_on.PST.PFV.3SG.N that company.SG.NOM
nie miała zgody.
 NEG have.PST.3SG.F permission.SG.GEN
 ‘The fact that the company did not have a permission escalated the problem.’
- c. *[Artykuły] nie zmieniłyby nic, a jedynie*
 [article.PL.NOM] NEG change.COND.3PL nothing but only
dodałyby oliwy do ognia.
 add.COND.3PL olive.SG.GEN to fire.SG.GEN
 ‘[The articles] would not change anything but they would only escalate the problem.’

⁵ *Wyłożyć kawę na ławę*: literally, ‘to lay the coffee on the table’. The figurative meaning of this idiom involves talking directly, usually about something unpleasant.

Test 8: DIFFERENT TENSE(S)

The test was conducted in order to establish whether an idiom may appear in various tenses.

- (17) *Żaden z nas nie puścił / puszcza /*
none from 1PL.GEN NEG let.PST.PFV.3SG.M let.PRS.IPFV.3SG
*puści pary z ust*⁶.
let.PRS.PFV.3SG steam.SG.GEN from mouth.PL.GEN
'None of us said / says / will say anything (about the issue).'

Test 9: DIFFERENT ASPECT(S)

The aim of the test was to verify whether an idiom may appear in various aspects. For instance, an idiom *dolewać oliwy do ognia* ('to add fuel to the fire') appeared both in perfective (*dolać*) as well as imperfective (*dolewać*) aspects.

- (18) a. *To tylko dolało oliwy do ognia.*
it only pour.PST.PFV.3SG.N olive.SG.GEN to fire.SG.GEN
'It only escalated the problem.'
- b. *Czy nie dolewam oliwy do ognia?*
Q NEG pour.PRS.IPFV.1SG olive.SG.GEN to fire.SG.GEN
'Do I escalate the problem?'

Test 10: PRONOMINAL COMPLEMENT'S SUBSTITUTION

The test was focused on the question whether a complement DP can be picked up by a pronoun. None of the selected idioms passed this test.

- (19) **Janek schował głowę w piasek i*
John.NOM hide.PST.PFV.3SG.M head.SG.ACC in sand.SG.ACC and
Maria też ją schowała w piasek.
Mary.NOM too 3SG.F.ACC hide.PST.PFV.3SG.F in sand.SG.ACC
'John avoided the responsibility and Mary avoided it, too.'

Test 11: NUMBER MODIFICATION

The aim of the test was to check whether the number of the complement DP could be changed.

⁶ *Puścić parę z ust*: literally, 'to let the steam go out of one's mouth'. The idiom could be translated as: "to let the cat out of the bag".

- (20) *Kto ukrywa prawdę? Czy ci, którzy zobowiązani prawem trzymają języki za zębami?*
 who.NOM hide.PRS.IPFV.3SG truth.SG.ACC Q these.PL.NOM who.PL.NOM
 obliged.PL.NOM law.SG.INS keep.PRS.IPFV.3PL tongue.PL.ACC behind tooth.PL.INS
 ‘Who keeps secrets? The ones who remain silent in the name of the law?’

Test 12: RELATIVIZATION

The test’s aim was to determine whether a complement might be followed by a relative clause.

- (21) *Maria nie wiedziała o istnieniu jeszcze jednego asa, którego Jan miał w rękawie.*
 Mary.NOM NEG know.PST.IPFV.3SG.F about existence.SG.LOC yet
 one.SG.GEN ace.SG.GEN which.SG.ACC John.NOM have.PST.3SG.M in
 sleeve.SG.LOC
 ‘Mary did not know about one more secret that will give John advantage.’

Test 13: LEFT DISLOCATION

The goal of the test was to verify whether a complement may be left-dislocated. Left dislocation is a syntactic movement operation in which a constituent (an argument or an adjunct of the clause) appears outside the left clause boundary.

None of the selected idioms passed this test without losing its figurative meaning.

- (22) a. **Byka, to chwycilem (go) za rogi.*
 bull.SG.ACC so grasp.PST.PFV.1SG.M 3SG.M.ACC by horn.PL.ACC
 (intended: ‘The opportunity, I took advantage of it.’)
- b. **Co do byka, to chwycilem go za rogi.*
 what to bull.SG.GEN so grasp.PST.PFV.1SG.M 3SG.M.ACC by
 horn.PL.ACC
 (intended: ‘When it comes to the opportunity, I took advantage of it.’)

On the basis of these tests, syntactic flexibility of 33 idioms was estimated and expressed on a scale from 0 to 13 (where “0” means “no test passed” and “13” means “all the tests passed”). The least syntactically flexible idioms in our study passed three of the 13 tests, and the most syntactically flexible idiom passed 10 of the 13 tests. The results of the tests are summarized in the Appendix. One conclusion which could be drawn at this point concerns syntactic flexibility as a scalar

quantity, rather than a zero-one property (see Fraser 1970; Gibbs and Gonzales 1985; Reagan 1987 for a similar observation), as illustrated for selected idioms in Figure 4.

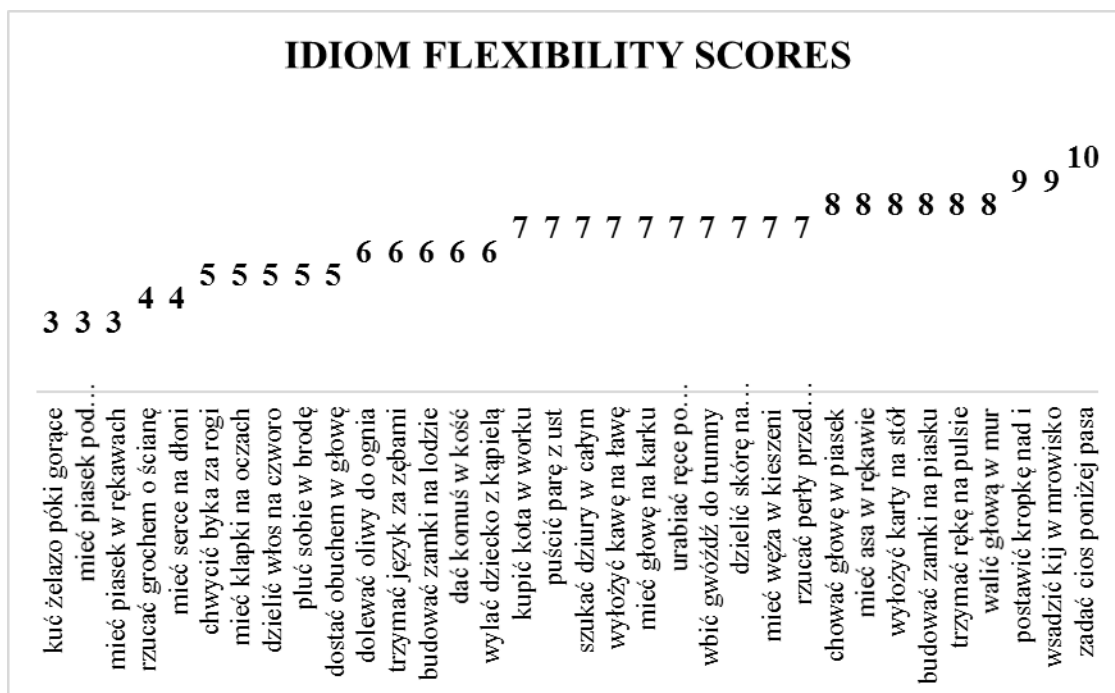
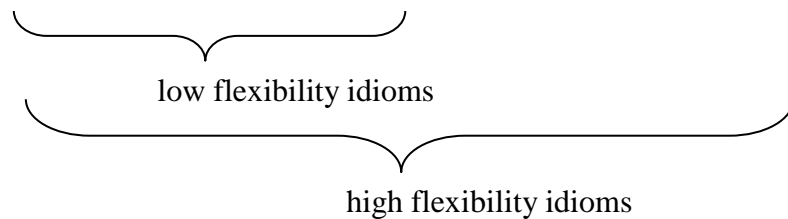


Figure 4. The results of syntactic flexibility tests

Interestingly, the kind of tests passed by the idiom groups does not seem to be random. While the majority of idioms, regardless of the syntactic flexibility score, allowed for verb ellipsis and verb substitution, the least syntactically flexible idioms (scores ‘3’, ‘4’, and ‘5’ in the classification in Figure 4) allowed only for most external modifications (at or above the TP level), i.e. different tense, modality, negation and external adverbial modification. More flexible idioms (score ‘6’), apart from the external modifications mentioned above, seemed to be much more susceptible to aspect modification and passive formation. Only most syntactically flexible idioms (score ‘7’ or above) allowed for VP-internal modifications, i.e., adjectival modification, number modification of the object noun or relativization. This finding suggests that idioms’ syntactic properties seem to reflect the hierarchy of projections proposed by major generative accounts (see, among others, Adger 2003; Cinque 1999; Ramchand and Svenonius 2014), with low flexible idioms only allowing for most external modifications related to higher functional projections (i.e., those above AspP), and high flexible idioms also allowing for modifications related to lower functional projections (i.e., those including AspP and VoiceP/vP) as well as modifications of elements within the VP, as depicted in (21). This resulted in a division of the tested idioms into two groups: low syntactic flexibility idioms (i.e., those with flexibility scores 3-5; 10 idioms altogether) and high syntactic

flexibility idioms (i.e., those with flexibility scores 6-10; 23 idioms in total) (see the Appendix for details).

(23) **Hierarchy of Projections:** Fin(itness)P > T(ense)P > (NegP) > Asp(ect)P > VoiceP /vP > VP



2.3 Experimental method

A cloze probability test with a response time measurement based on Staub et al. (2015) was conducted. This method was selected due to its advantages over its “traditional” paper-and-pencil equivalent in that, instead of writing down their responses, the participants provide oral continuations of expressions presented to them and their answers are recorded. Consequently, it is possible to measure not only idiom cloze probabilities but also idiom recognition reaction times. In the context of idiom-related studies, cloze probability could be defined as a number of expected (i.e., idiomatic) continuations provided by the participants and the reaction (response) time refers to the interval between the prompt and the onset of the response.

2.4 Participants

20 university students (14 females, mean age: 19.7, age range from 19 to 23 years old), all of whom were native speakers of Polish. All reported normal or corrected-to-normal vision and participated in the study for a course credit.

2.5 Experimental design

Based on the tests described in section 2.2, the idioms were divided in two types depending on their SYNTACTIC FLEXIBILITY: high flexibility and low flexibility. In addition, all idioms’ frequencies were checked in Poliqarp 1.0 search engine for the National Corpus of Polish (NKJP) (Janus and Przepiórkowski 2007). The selected high flexibility and low flexibility idioms did not differ significantly in terms of their lemma frequency ($W = 118$, $p\text{-value} = 0.10$).

Each of the 33 idioms were presented in three different PROMPT LOCATIONS (see section 2.6), as depicted in Table 1. For each prompt location, cloze probability (CP) and response times (RT) were measured.

Table 1. Experimental design

	PROMPT LOCATION 1	PROMPT LOCATION 2	PROMPT LOCATION 3
HIGH FLEXIBILITY	<i>verb</i>	<i>verb + DP_{object}</i>	<i>verb + DP_{object} + preposition</i>
LOW FLEXIBILITY	<i>verb</i>	<i>verb + DP_{object}</i>	<i>verb + DP_{object} + preposition</i>

All idioms were presented in an infinitive form and imperfective aspect, as illustrated in (22):

- (24) a. **PROMPT LOCATION 1: “VERB”:**
dolewać
 add
- b. **PROMPT LOCATION 2: “VERB + DP_{object}”**
dolewać oliwy
 add fuel
- c. **PROMPT LOCATION 3: “VERB + DP_{object} + PREPOSTION”**
dolewać oliwy do
 add fuel to

Additionally, 153 fillers (literal collocations and figurative expressions other than idioms, such as similes) were presented to the participants. Idiom triplets (and fillers) were equally distributed across three experimental lists. This was necessary to avoid the situation when one participant sees the same idiom in all the three PROMPT LOCATIONS. In other words, subjects who saw one idiom in PROMPT LOCATION 1 did *not* see it in PROMPT LOCATION 2 and PROMPT LOCATION 3. Latin-square design was used in order to ensure the equal distribution of items across the lists. Each list was presented to an independent group of participants.

2.6 Procedure

Participants were tested individually in a quiet room. They sat comfortably in front of a Windows PC computer. Their responses were recorded by the computer’s built-in microphone. PsychoPy software, version 1.83.04 (Peirce 2007, 2009) was used to present the experimental stimuli. An experimental session lasted about 20 minutes.

An experimental session consisted of a training session and four experimental blocks divided by optional breaks. The participants were instructed to say aloud the first word (or words) which they could think of as a completion of an expression which was presented to them. All expressions were displayed in the same way (see Figure 5): the word(s) an expression began with was (were) visible on the screen for 300 ms each, followed by a horizontal line which served as a prompt to provide an oral continuation. A fixation cross appeared on the screen at the beginning of each new trial.

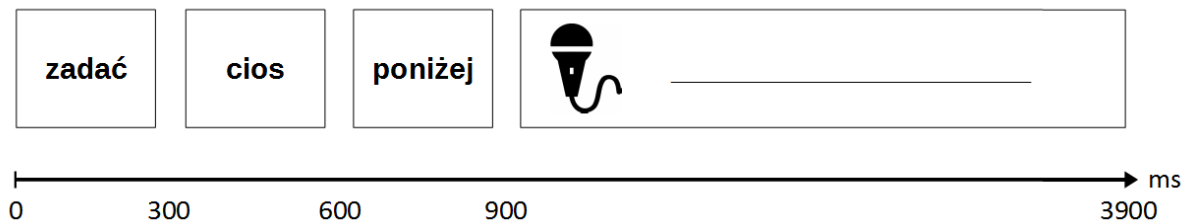


Figure 5. Experimental procedure

During the trials, the PsychoPy (1.83.04) software generated .wav files containing the subjects' oral responses. On the basis of the obtained data, transcriptions of all the responses were prepared and, as a consequence, the idiom cloze probabilities (CP) (the number of idiomatic completions) were calculated. Also, idiom response times (RT) were estimated: mean response latencies (for idiomatic completions only) were calculated on the basis of waveforms generated and interpreted with the aid of the Praat software (see Staub et al. 2015).

2.7 Predictions

We predicted that irrespective of the syntactic flexibility, for all idioms we should observe increasing cloze probabilities and decreasing response times for successive PROMPT LOCATIONS. However, as has been explained in section 1, Nunberg et al.'s (1994) and Sprenger et al.'s (2006) approaches lead to competing predictions for the comparisons of the cloze probabilities and response times for low and high flexibility idioms.

Prediction 1 (based on Nunberg et al. 1994): Low flexibility idioms (idiomatic phrases) should give rise to higher CP rates (i.e., more idioms should be recognized in a given prompt location) and faster RTs (i.e., idiomatic continuations of the recognized idioms should be produced faster) as compared to high flexibility idioms (idiomatic expressions).

Prediction 2 (based on Sprenger et al. 2006): Syntactic flexibility should not be a factor influencing the idiom recognition. No significant differences in cloze probability rates and response times between idioms depending on their degree of syntactic flexibility are expected.

3 Results

Statistical data analysis was carried out using R software (R Development Core Team 2017). The analysis was carried out on complete observations only. As a complete set of observations was not always recorded for each of 33 idioms with respect to a given participant, the resulting data set varied in the number of observations for prompt location and syntactic flexibility.

RTs and CPs were measured for: (i) PROMPT LOCATION (3 levels: after the verb, after the object, and after the preposition) and (ii) SYNTACTIC FLEXIBILITY of the idiom (2 levels: low flexibility (scores from 3 to 5), and high flexibility (scores from 6 to 10)).

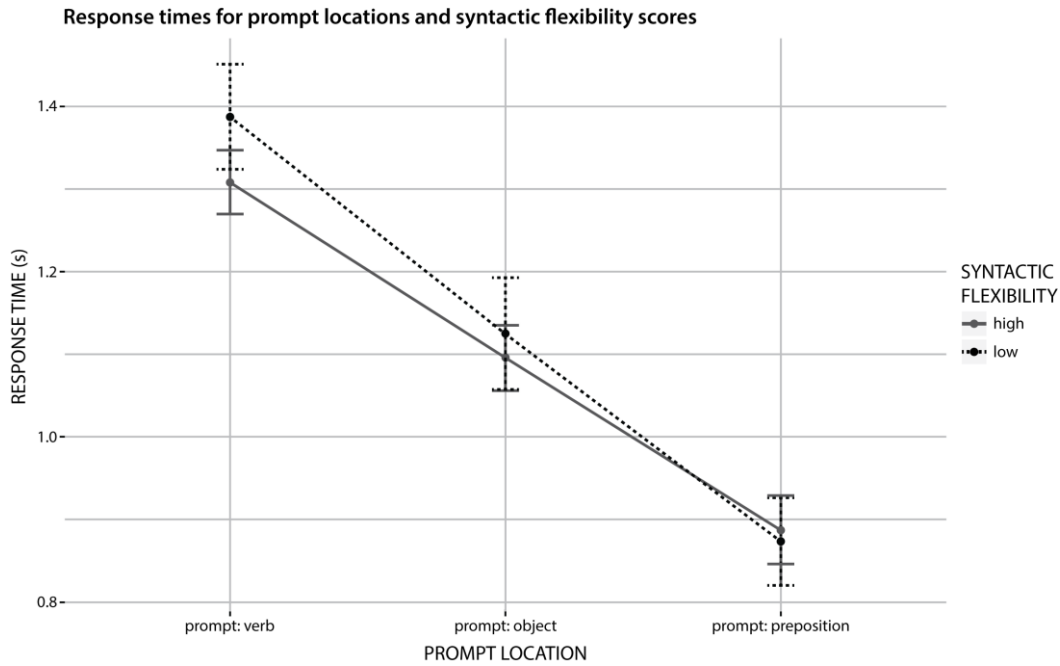
To determine whether the PROMPT LOCATION or SYNTACTIC FLEXIBILITY of the idiom influence the time needed to recognize the idiom, a generalized linear model analysis using the *lme* function from *nlme* (Pinheiro et al. 2017) package was fitted. Participants and items (idioms) were set as random effects. The significance of main effects and interaction effect was estimated with log-likelihood on model comparison. For main effects we compared a model with one effect to a model with two effects. As for an interaction effect, we compared the model with two main effects to a model containing their interaction. As for the cloze probability, since the response variable was a binary one, a loglinear analysis using the *loglm* function from *MASS* (Ripley et al. 2017) package was carried out. Parallel to the model for response times, participants and items (idioms) were set as random effects. The complete model consisted of frequencies for the correct and incorrect idiom recognition distributed across prompt locations and idiom syntactic flexibility groups. The significance of main effects and interaction effects was estimated with log-likelihood on model comparison. Significance of each interaction and main effect was determined by subtracting the effect of interest from a model and comparing this model to the model with all main and interaction effects.

3.1 Results for response times

The descriptive statistics of idiom production time analysis is summarized in Table 2, which presents mean response times for high and low FLEXIBILITY idioms presented in three PROMPT LOCATION conditions. Mean response times for the two idiom groups are presented graphically in Figure 6.

Table 2. Descriptive statistics

PROMPT LOCATION	SYNTACTIC FLEXIBILITY									
	high					low				
	N	mean RT	SD	SE	CI	N	mean RT	SD	SE	CI
verb	135	1.308	0.468	0.040	0.080	60	1.387	0.493	0.064	0.127
object	135	1.085	0.468	0.040	0.080	65	1.125	0.546	0.068	0.135
preposition	145	0.863	0.484	0.040	0.080	60	0.874	0.412	0.053	0.106

**Figure 6.** Mean response times

Statistical analysis of response times showed a significant main effect for a prompt location only ($\chi^2(1) = 167.01$ $p < .0001$). When the effect was broken down to pairwise comparisons, it revealed that response times differ significantly between all the three PROMPT LOCATIONS; see Table 3.

Table 3. Significant differences in RTs

PROMPT LOCATION	t-value	p-value
verb vs. object	19.68	< .0001
object vs. preposition	8.21	< .0001
verb vs. preposition	14.21	< .0001

These results, combined with the visual inspection of the mean RTs, indicate that the idioms are recognized significantly faster with the increase of the idiom chunk the participants are presented

with, regardless of the syntactic flexibility of the idiom. In other words, the longer the stimulus was, the faster it was completed by the participants.

Crucially for our predictions (see section 2.7), no significant results were found for the main effect of SYNTACTIC FLEXIBILITY ($\chi^2(1) = 0.854$, $p = 0.35$) or the interaction between PROMPT LOCATION and SYNTACTIC FLEXIBILITY ($\chi^2(1) = 0.188$, $p = 0.91$). Also, a by-subject analysis, whose aim was to take into account potential individual differences between the participants, revealed no significant interaction effect between the participant and idiom SYNTACTIC FLEXIBILITY ($p = 0.0976$), which suggests that the individual subjects' RTs did not differ significantly with respect to the division between high and low flexibility idioms. This would additionally confirm that the differences between the two idiom groups as to their degree of SYNTACTIC FLEXIBILITY did not influence idiom production times.

3.2 Results for cloze probability

The numbers for recognized and not recognized idioms of high and low SYNTACTIC FLEXIBILITY displayed in three PROMPT LOCATIONS are presented below in Table 4.

Table 4. The numbers of recognized vs. not recognized idioms

PROMPT LOCATION	SYNTACTIC FLEXIBILITY			
	high		low	
	no. recognized	no. not recognized	no. recognized	no. not recognized
verb	10	125	2	58
object	83	52	35	30
preposition	110	35	43	17

The three-way loglinear analysis (PROMPT LOCATION x SYNTACTIC FLEXIBILITY x CLOZE PROBABILITY) produced the final model that retained only one two-way interaction effect. The likelihood of this model was ($\chi^2(0) = 0$, $p = 1$). This indicated that a two-way interaction between PROMPT LOCATION and CLOZE PROBABILITY ($\chi^2(2) = 231.99$, $p < .001$) is significant. Accordingly, PROMPT LOCATION can be treated as a substantial factor determining idiom recognition or lack of it. This effect is presented graphically below in Figure 7.

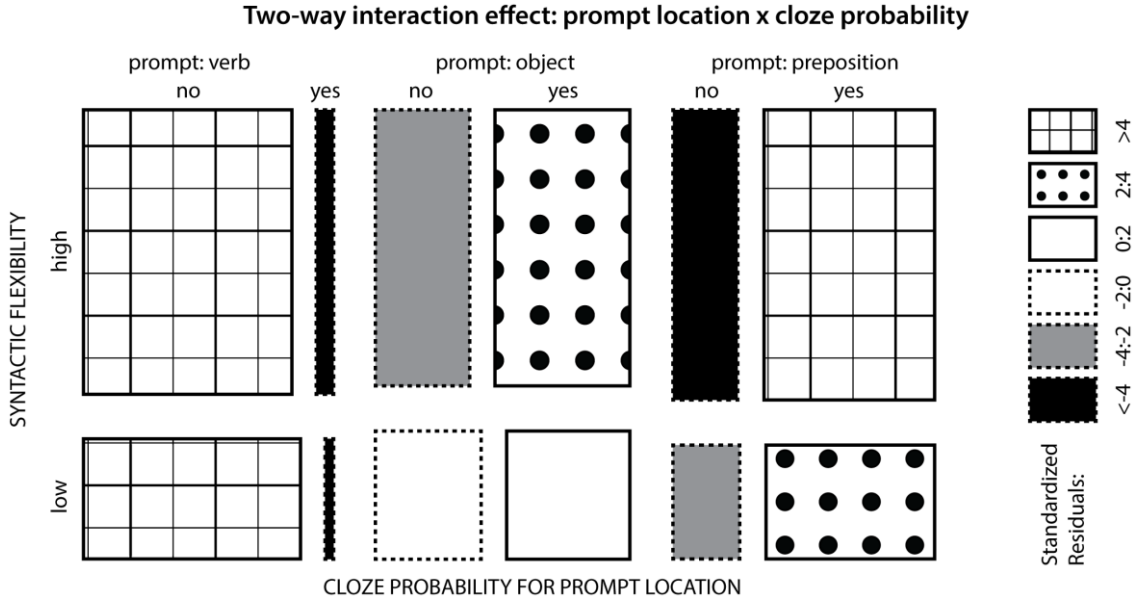


Figure 7. Prompt location x cloze probability

The differences in frequencies for idiom recognition and lack of it between PROMPT LOCATIONS are significant ($\chi^2(2) = 204.93$, $p < .001$). Therefore, the revealed effect can be interpreted as follows: regardless of the SYNTACTIC FLEXIBILITY of the idiom, when the prompt was located after the verb, idiom recognition was significantly hindered. When the prompt was located after the object or the preposition, idiom recognition was significantly improved. Importantly for our predictions (see section 2.7), however, no statistically-significant difference was observed with respect to idioms' SYNTACTIC FLEXIBILITY. This would suggest that SYNTACTIC FLEXIBILITY should not be considered a factor influencing idiom recognition.

4 Conclusion, discussion and open questions

The statistical analysis confirmed that irrespective of the SYNTACTIC FLEXIBILITY, for all idioms, cloze probabilities increased and response times decreased for successive PROMPT LOCATIONS. Furthermore, the obtained results suggest that syntactic flexibility does not affect idiom recognition. We did not observe any significant difference in cloze probability rates and response time measures between high flexibility and low flexibility idioms. This result is compatible with **Prediction 2**, which was formulated on the basis of Sprenger et al.'s (2006) "superlemma theory". According to this model, all idioms have the same hybrid representation in the mental lexicon containing information about the idiom's figurative meaning, its constituent lemmas and syntactic idiosyncrasies (recall Figures 2 and 3). The superlemma representation is accessed after the comprehenders have reached Cacciari and Tabossi's (1988) idiom recognition point, that is, after they have recognized that the processed string of words, in fact, forms an idiomatic configuration. Given that the information about the idiom's syntactic flexibility is

accessible only after one reaches the idiom's superlemma representation, this factor cannot facilitate idiom recognition. However, one must be cautious in generalizing our results to natural production. In spite of the fact that in our experiment the participants were asked to produce idiom continuations, this was not a standard production experiment. Rather, it combined the elements of reception and production tasks. As we were interested in the influence of syntactic flexibility on how fast idioms are retrieved from the mental lexicon, the subjects were initially presented with, and hence expected to process, the experimental stimuli (i.e., a fragment of an idiomatic phrase which was to prompt their recognition of a given idiom). Only having recognized an idiom, the participants could produce its continuation. That is why, the part of the experiment which involved actual production took place after an idiom had already been recognized and after the superlemma had been accessed. From the speaker's perspective then, our study differed significantly from the standard production tasks (such as, for instance, picture description) in which a full idiomatic representation is accessible already at the conceptual level, and only then the message is mapped onto the syntactic and semantic structure of a given language, as pointed out in Sprenger et al. (2006: 163). In ordinary idiom production, an idiomatic utterance that can be taken either literally or not is then an outcome rather than a stimulus. In other words, the speakers have access to the idiom's superlemma representation (including, among other things, the information about its syntactic flexibility) pre-verbally, that is, before they start producing the idiomatic string of words. It is not excluded that in this case syntactic flexibility plays a role. Another open question is why idioms differ with respect to the degree of their syntactic flexibility (see Wierzba 2016 for a discussion on this topic for German idioms). If, as suggested in Sprenger et al.'s superlemma model, the idiom's superlemma representation contains information about its idiosyncratic syntactic properties, it is still unclear how exactly the lexical representation of an idiom is mapped onto syntax.

Another aspect in need for further research is connected to Sprenger et al.'s (2006: 176) observation that the selection and processing of the superlemma are the same as the selection and processing of single words. If so, one could assume that factors such as context, frequency and familiarity may be relevant for the theory of how idioms are accessed from the mental lexicon (see Marelli and Luzzatti 2012; White 2008, among others, for discussion).

References

- Adger, D. 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.
- Benko, V. 2014. "Aranea: Yet another family of (comparable) web corpora." In: P. Sojka, A. Horák, I. Kopeček, and K. Pala (eds.) *TSD 2014, LNAI 8655*. Cham, Switzerland: Springer International Publishing. 257-264.
- Bobrow, S. A., and Bell, S. M. 1973. "On catching on to idiomatic expressions." *Memory and Cognition* 1. 343-346.
- Cacciari, C., and Glucksberg, S. 1991. "Understanding idiomatic expressions: The contribution of word meanings." In: G. Simpson (ed.) *Understanding Word and Sentence*. Amsterdam, Netherlands: North-Holland. 217-240.
- Cacciari, C., and Tabossi, P. 1988. "The comprehension of idioms." *Journal of Memory and Language* 27. 668-683.
- Chomsky, N. 1980. *Rules and Representations*. New York: Columbia University Press.
- Cieślicka, A. 2012. "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(6). 475-503.
- Cinque, G. 1999. *Adverbs and Functional Heads. A Cross-linguistic Perspective*. New York: Oxford University Press.
- Cutting, J. C., and Bock, K. 1997. "That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends." *Memory and Cognition* 25(1). 57-71.
- Fraser, B. 1970. "Idioms within a transformational grammar." *Foundations of Language* 6. 22-42.
- Geeraerts, D. 1995. "Specialization and reinterpretation in idioms." In: M. Everaert, E.-J. van der Linden, A. Schenk, and R. Schreuder (eds.) *Idioms: Structural and Psychological Perspectives*. Hillsdale, NJ: Lawrence Erlbaum Associates. 57-74.
- Gibbs, R. W., and Gonzales, G. P. 1985. "Syntactic frozenness in processing and remembering idioms." *Cognition* 20(3). 243-259.
- Gibbs, R. W., and Nayak, N. P. 1989. "Psycholinguistic studies on the syntactic behavior of idioms." *Cognitive Psychology* 21. 100-138.
- Gibbs, R. W., Nayak, N., Bolton, J., and Keppel, M. 1989a. "Speakers' assumptions about the lexical flexibility of idioms." *Memory and Cognition* 17. 58-68.
- Gibbs, R. W., Nayak, N. P., and Cutting, C. 1989b. "How to kick the bucket and not decompose: Analyzability and idiom processing." *Journal of Memory and Language* 28. 576-593.
- Glasbey, S. R. 2003. "Let's paint the town red for a few hours: Composition of aspect in idioms." In: A. M. Wallington (ed.) *Proceedings of the ACL workshop on Lexicon and Figurative Language*. Stroudsburg, PA: Association for Computational Linguistics. 42-48.

- Janus, D. and Przepiórkowski, A. 2007. "Poliquarp 1.0: Some technical aspects of a linguistic search engine for large corpora." In: J. Waliński, K. Kredens, and S. Goźdz-Roszkowski (eds.) *The Proceedings of Practical Applications in Language and Computers PALC 2005*. Frankfurt am Main: Peter Lang.
- Katz, J. 1973. "Compositionality, idiomaticity, and lexical substitution." In: S. R. Anderson and P. Kiparsky (eds.) *A Festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston. 357-76.
- Katz, J., and Postal, P. 1963. "Semantic interpretation of idioms and sentences containing them." *Quarterly Progress Report of the MIT Research Laboratory of Electronics* 70. 275-82.
- Kędzierska, H. 2018. "The role of context in the processing of idioms: An experimental study." *Generative Linguistics in Wrocław (GLiW)*, vol. 5. Wrocław: Center for General and Comparative Linguistics (CGCL).
- Leivada, E. 2017. "The primitives of the lexicon: Insights from aspect in idioms." *Glossa: A Journal of General Linguistics* 2(1): 105. 1-23.
- Machonis, P. A. 1985. "Transformations of verb phrase idioms: Passivization, particle movement, dative shift." *American Speech* 60. 291-308.
- Marelli M., and Luzzatti C. 2012. "Frequency effects in the processing of Italian nominal compounds: Modulation of headedness and semantic transparency." *Journal of Memory and Language* 66(4). 644-664.
- Müller, G. 2000. "Idioms and transformations." A talk given at the GGS meeting in Potsdam.
- Nunberg, G. 1979. "The non-uniqueness of semantic solutions: Polysemy." *Linguistics and Philosophy* 3(2). 143-184.
- Nunberg, G., Sag, I., and Wasow, T. 1994. "Idioms." *Language* 70(3). 491-538.
- Peirce, J. W. 2007. "PsychoPy – Psychophysics software in Python." *Journal of Neuroscience Methods* 162(1-2). 8-13.
- Peirce, J. W. 2009. "Generating stimuli for neuroscience using PsychoPy." *Frontiers in Neuroinformatics* 2(10).
- Peterson, R. R., Burgess, C., Dell, G., and Eberhard, K. M. 2001. "Dissociation between syntactic and semantic processing during idiom comprehension." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27. 1223-1237.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., EISPAC, Heisterkamp, S., and van Willigen, B., R-core. 2017. "Package 'nlme': Linear and Nonlinear Mixed Effects Models."
- Ramchand, G. C., and Svenonius, P. 2014. "Deriving the functional hierarchy." *Language Sciences* 46. 152-174.

- R Core Team. 2017. "R: A language and environment for statistical computing. R Foundation for Statistical Computing." Vienna, Austria.
- Reagan, R. T. 1987. "The syntax of English idioms: Can the dog be put on?" *Journal of Psycholinguistic Research* 16. 417-441.
- Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., and Firth, D. 2017. MASS: Functions and datasets to support Venables and Ripley, "Modern Applied Statistics with S" (4th edition, 2002).
- Sprenger, S. A., Levelt, W. J. M., and Kampen, G. 2006. "Lexical access during the production of idiomatic phrases." *Journal of Memory and Language* 54(2). 161-184.
- Staub, A., Grant, M., Astheimer, L., and Cohen, A. 2015. "The influence of cloze probability and item constraint on close task response time." *Journal of Memory and Language* 82. 1-17.
- Stone, M. 2016. *The Difference between Bucket-Kicking and Kicking the Bucket: Understanding Idiom Flexibility*. PhD thesis, The University of Arizona.
- Swinney, D. A., and Cutler, A. 1979. "The access and processing of idiomatic expressions." *Journal of Verbal Learning & Verbal Behavior* 18. 523-534.
- Vespignani, F., Canal, P., Molinaro, N., Fonda, S., and Cacciari, C. 2010. "Predictive mechanisms in idiom comprehension." *Journal of Cognitive Neuroscience* 22(8). 1682-1700.
- Wasow, T., Sag, I., and Nunberg, G. 1983. "Idioms: An interim report." In: S. Hattori and K. Inoue (eds.) *Proceedings of the XIIIth International Congress of Linguistics* Tokyo: CIP. 102-105.
- White, S. J. 2008. "Eye-movement control during reading: Effects of word frequency and orthographic familiarity." *Journal of Experimental Psychology Human Perception and Performance* 34(1). 205-223.
- Wierzba, M. 2016. "An experimental view on the syntactic flexibility of German idioms." A talk given at Budapest-Potsdam-Lund Colloquium 2016.

APPENDIX

IDIOM	FLEX (SCORE)	RT ⁷ verb	RT obj.	RT prep.	CP ⁸ verb	CP obj.	CP prep.
<i>budować zamki na lodzie</i> (‘to build castles on the ice’)	high (6)	1.036857	1.0905	1.131167	0	0	0
<i>budować zamki na piasku</i> (‘to build castles on the sand’)	high (8)	1.036857	1.1982	1.084286	0	0	5
<i>chować głowę w piasek</i> (‘to hide your head in the sand’)	high (8)	1.269143	0.9585	1.009095	2	6	6
<i>chwycić byka za rogi</i> (‘to take the bull by the horns’)	low (5)	1.5382	0.717	0.629429	0	7	6
<i>dolewać oliwy do ognia</i> (‘to add olive to the fire’)	high (6)	1.0216	0.712119	0.5725	1	7	6
<i>dostać obuchem w głowę</i> (‘to be hit by a war-hammer’)	low (5)	1.2365	1.398029	1.012286	0	4	4
<i>dzielić skórę na niedźwiedziu</i> (‘to share the skin on a bear’)	high (7)	1.2854	0.9854	1.079167	0	3	3
<i>dzielić włos na czworo</i> (‘to split a hair into four parts’)	low (5)	1.2854	1.135714	0.735333	0	3	6
<i>kuć żelazo póki gorące</i> (‘to forge the iron while it’s hot’)	low (3)	1.050571	0.937167	0.615	2	6	5
<i>kupować kota w worku</i> (‘to buy a cat in a sack’)	high (7)	1.905333	1.071167	0.734716	0	5	6
<i>mieć asa w rękawie</i> (‘to have an ace in a sleeve’)	high (8)	1.507167	0.8155	0.613167	0	6	6
<i>mieć głowę na karku</i> (‘to have a head on the neck’)	high (7)	1.507167	1.038	0.679	0	4	6
<i>mieć klapki na oczach</i> (‘to have blinkers on the eyes’)	low (5)	1.507167	0.883833	0.645143	0	5	6
<i>mieć piasek pod powiekami</i> (‘to have sand under one’s eyelids’)	low (3)	1.5322	1.485333	1.474349	0	0	1
<i>mieć piasek w rękawach</i> (‘to have sand in one’s sleeves’)	low (3)	1.507167	1.485333	1.387333	0	0	0
<i>mieć serce na dłoni</i> (‘to have a heart on one’s hand’)	low (4)	1.507167	1.256556	0.9384	0	0	5
<i>mieć węża w kieszeni</i>	high (7)	1.507167	1.1708	1.119	0	3	6

⁷ Abbreviations “RT verb / object / preposition” refer to the mean response times elicited when the prompt was located after the verb / object / preposition.

⁸ Similarly, abbreviations “CP verb / object / preposition” refer to the mean cloze probabilities elicited when the prompt was located after the verb / object / preposition.

(‘to have a snake in a pocket’)							
<i>pluć sobie w brodę</i> (‘to split at one’s own chin’)	low (5)	1.455714	1.146286	0.5785	0	4	2
<i>postawić kropkę nad i</i> (‘to put a dot above an i’)	high (9)	1.218143	0.800833	0.68772	0	5	7
<i>puścić parę z ust</i> (‘to let the steam go one’s mouth’)	high (7)	1.18875	1.018714	0.8205	0	5	5
<i>rzucić grochem o ścianę</i> (‘to throw the beans against a wall’)	low (4)	1.331333	0.863813	0.677429	0	6	7
<i>rzucić perły przed wieprze</i> (‘to throw the pearls in front of the pigs’)	high (7)	1.331333	1.713833	1.5706	0	0	0
<i>szukać dziury w całym</i> (‘to look for a hole in the whole’)	high (7)	1.029251	0.858143	0.619333	1	6	6
<i>trzymać język za zębami</i> (‘to hold one’s tongue behind teeth’)	high (6)	1.4905	0.841368	0.542714	1	6	7
<i>trzymać rękę na pulsie</i> (‘to hold your hand on the pulse’)	high (8)	1.4905	1.4304	0.784429	0	2	6
<i>urabiać ręce po łokcie</i> (‘to mould one’s hands up to elbows’)	high (7)	1.47566	1.2145	0.6632	0	3	5
<i>walić głową w mur</i> (‘to bang one’s head again a wall’)	high (8)	1.4335	1.031667	0.693857	2	4	3
<i>wbić gwóźdź do trumny</i> (‘to drive the nail into the coffin’)	high (7)	1.199429	1.2855	1.131	0	1	2
<i>wsadzić kij w mrowisko</i> (‘to put a stick into an anthill’)	high (9)	1.42925	1.04525	0.834	0	4	5
<i>wylać dziecko z kąpielą</i> (‘to throw a baby with the bathwater’)	high (6)	1.02825	1.8576	1.633667	0	2	1
<i>wyłożyć karty na stół</i> (‘to lay one’s cards on the table’)	high (8)	1.457	0.943571	0.738429	1	5	6
<i>wyłożyć kawę na ławę</i> (‘to put the coffee on the bench’)	high (7)	1.457	0.880251	0.594286	2	6	7
<i>zadać cios poniżej pasa</i> (‘to punch somebody below the belt’)	high (10)	0.810114	1.335	0.667166	0	1	7