

## RESEARCH

# The acquisition of Hebrew idioms: Stages, internal composition, and implications for storage

Julie Fadlon<sup>1</sup>, Julia Horvath<sup>2</sup>, Tal Siloni<sup>2</sup> and Kenneth Wexler<sup>3</sup>

<sup>1</sup> The Hebrew University of Jerusalem, Mount Scopus, 91905, Jerusalem, IL

<sup>2</sup> Tel-Aviv University, Ramat Aviv, 69978, Tel-Aviv, IL

<sup>3</sup> MIT, 77 Massachusetts Ave, Cambridge, MA 02139, US

Corresponding author: Julie Fadlon ([Julia.Fadlon@mail.huji.ac.il](mailto:Julia.Fadlon@mail.huji.ac.il))

---

The study investigates the potential effects of the internal structure of idioms on their acquisition. It tested school-children (1<sup>st</sup> to 3<sup>rd</sup> graders) acquiring Hebrew. Comprehension and production experiments examined the effect of two structural factors on the acquisition of verb phrase idioms: (i) whether the idiom was a full lexically fixed constituent or involved an open slot, namely a free, lexically unspecified obligatory constituent; (ii) whether or not the idiom was decomposable. While neither (i) nor (ii) influenced idiom comprehension in these age groups, idiom production was affected by both. In the production experiment, performance with nondecomposable idioms was significantly better than performance with decomposable idioms across age groups. Further, an analysis by age group showed significant interactions of factors (i) and (ii) for second and third graders. We propose that the main effect of (non)decomposability is due to two distinct techniques (available in grammar) that children utilize for the storage of idioms, and to children's facility with retrieval of units vs. retrieval by composition. Children, unlike adults, store nondecomposable phrasal idioms as independent entries, rather than as subentries of their lexical head. The reason for this misanalysis, we propose, is that children have difficulty reconciling the constituent structure of nondecomposable idioms with their lack of semantic composition. The effect of an open slot differs in accordance with the storage technique: It facilitates retrieval of units because there are fewer lexically fixed constituents to recover, but makes retrieval of subentries harder due to the nonuniform lexical representation of the idiom.

---

**Keywords:** idiom acquisition; Hebrew; idiom representation; language acquisition; production; comprehension

---

## 1 Setting the stage

Idioms exhibit an internal duality. On the one hand, they are complex entities whose internal makeup reflects structural properties of phrasal units. But on the other hand, they have conventionalized meaning that cannot be predicted based on the meaning of their building blocks and their structural properties. Therefore, the question as to how they are acquired is of particular interest.

Previous studies of first language (L1) acquisition of idioms have focused mainly on Italian, English, and French, examining (a) the relationship between the maturation of figurative language and the acquisition of idioms by children (Levorato & Cacciari 1992; Levorato, Nesi & Cacciari 1995; 2004; Laval 2003); (b) the use of an informative context in the comprehension and production of idioms (Gibbs 1987; Nippold & Tarrant Martin 1989; Laval 2003; Levorato & Cacciari 1995; 1999; Cain, Towse & Knight 2008); (c) the effect of idiom transparency/opacity and compositionality on the acquisition of idioms (Gibbs 1987; 1991; Nippold & Rudzinski 1993; Nippold & Taylor 1995; 2002; Levorato & Cacciari 1999; Caillies & Le Sourn-Bissaoui 2006; Cain, Towse & Knight 2008); and (d)

the developmental differences between idiom comprehension and production (Cacciari & Levorato 1989; Levorato & Cacciari 1992).

These studies have concluded that the maturation of figurative language is a necessary ingredient of children's achieving adult-like performance in comprehension and use of idioms, and that informative context and transparency facilitate these tasks. Further, it has been shown that the ability to understand and produce idioms develops gradually from age 6–7 and continues to mature even beyond the age of 12. Finally, it has been concluded that children are better at comprehending idioms than at producing them.

This study sets out to investigate the acquisition of verb phrase idioms along structural dimensions to be detailed in 1.1. Moreover, the study extends the range of languages explored regarding the process of idiom acquisition and its stages. It targets school-children acquiring Hebrew, a (non-Indo-European) language that has not yet been investigated in this respect. Finally, the study turns out to shed light on how idiomatic phrases are stored by children and in what way children's idiom storage and retrieval differs from the adult state.

### 1.1 Structural distinctions: Decomposability and open slots

The experiments we ran examined the effect of two structural factors on idiom acquisition: (i) whether the idiom was a full lexically fixed constituent or involved an open slot, namely a free, lexically unspecified but obligatory, subconstituent; (ii) whether the idiom was decomposable or nondecomposable (definition provided below).

The first factor concerns the fact that some VP-idioms include an open slot, and some VP-idioms do not. By open slot we mean an obligatory subconstituent of the idiom, part of its internal composition, which is not limited to a specific lexical choice, and therefore is filled by non-idiomatic lexical material in the course of the derivation (see also Mishani-Uval & Siloni 2016). Accordingly, we compared children's performance with verb phrase idioms that are full (lexically fixed) constituents to their performance with verb phrase idioms that include an obligatory open subconstituent to be filled by non-idiomatic material. The former will henceforth be referred to as *full idioms*, and the latter as *open-slot idioms*. Examples (1) and (2) illustrate this distinction (the open slot is marked by X).

- (1) Full idioms:
  - a. shoot the breeze
  - b. kill two birds with one stone
  
- (2) Open-slot idioms:
  - a. show X the door
  - b. take X to the cleaners

Factor (i) (full vs. open slot) can be expected to have an effect on the acquisition process. On the one hand, it is possible that children would show better knowledge of open-slot idioms, as these specify less fixed information that needs to be encoded and retrieved. By "less" we are referring to the amount of information that the child needs to learn and retrieve for a given idiom. When a slot within a VP idiom is an open one, namely, obligatory but not limited to a specific lexical choice, it has less information specified for it than a parallel slot with a fixed lexical choice specific to the idiom. Hence, open-slot idioms may be easier to learn and process. On the other hand, it is possible that the lack of uniformity in the type of information open-slot idioms encode, i.e. the mix of lexically fixed elements and open slots, would make them harder to learn and process. As far as we know, this distinction has not yet received attention in the acquisition literature.

The other factor, (ii), which we label *decomposability*, distinguishes between *decomposable idioms* vs. *nondecomposable idioms*. This distinction, in its current form, can be attributed to Nunberg, Sag & Wasow's (1994) seminal work on idioms. Their work clarified the distinction between these two types of idioms, which they labelled "idiomatically combining expressions" ("decomposable" idioms in our terminology) and idiomatic phrases ("nondecomposable" idioms), and proposed some potential syntactic correlates (specifically, flexibility, to be discussed briefly at the end of this section). The defining property of decomposable idioms was taken by Nunberg, Sag & Wasow to be the homomorphic mapping they exhibit between their literal and their idiomatic interpretation. Nondecomposable idioms show no such homomorphism. Van de Voort & Vonk (1994) discuss the same notion in terms of the "isomorphicity" of idioms, defined as the extent to which an idiom's metaphorical meaning can be distributed over its parts. The decomposability distinction itself, its defining criteria and correlates have been the subject of intense empirical scrutiny and debate over the past couple of decades in theoretical, corpus-based and psycholinguistic studies of idioms (for instance, Gibbs & Nayak 1989; Gibbs 1991; Titone & Connine 1994; Schenk 1995; Levorato & Cacciari 1999; Webelhuth & Ackerman 1999; Riehemann 2001; Caillies & Le Sourn-Bissaoui 2006; Wulff 2010).

Our criterion for classifying an idiom as decomposable is based on Nunberg, Sag & Wasow's and Van de Voort & Vonk's definitions. Our definition of decomposability is explicit regarding how we determine what the literature refers to as homomorphism or isomorphism. First, it requires that decomposability be determined based on a comparison of the idiom with the literal expression most closely matching its meaning. Second, the idiom's constituent structure is compared with the constituent structure of the literal expression, checking correspondence of constituent to constituent, where by constituent we mean the (lexical) head and the head's dependents, that is, its arguments and the adjuncts of its projection. Our definition is given in (3), and corresponding examples of decomposable vs. nondecomposable idioms in (4) and (5), respectively.

(3) Decomposability

An idiom is *decomposable* if it is isomorphic with its meaning in the sense that each of the idiom's constituents (the head and its dependents) corresponds to a constituent in the literal expression most precisely conveying the idiom's meaning; otherwise it is *nondecomposable*.

Accordingly, idioms like the ones in (4), for which there is isomorphism of constituent structure between the idiom and its interpretation as shown by the brackets, were considered *decomposable*, while idioms like the ones in (5), for which there is no such correspondence, were considered *nondecomposable*. The idioms in (4a–b) and (5a–b) are usually used in the literature to illustrate the distinction between decomposable vs. nondecomposable idioms, as their classification is straightforward. In addition to them, we also discuss a couple of prima facie less clear cases (in particular, (4c) and (5c)) below in order to be explicit as to how we applied the criterion in (3).

- (4) a. [spill] [the beans]  
       ['reveal] [the secret]'  
 b. [kill]<sub>1</sub> [two birds] [with one stone]  
       [achieve] [two goals] [with one action]  
 c. [take] [X] [with a grain of salt]  
       [view] [X] [with skepticism]

- (5) a. [kick] [the bucket]  
       ‘[die]’  
       b. [shoot] [the breeze]  
       ‘[chat]’  
       c. [put] [X’s cards] [on the table]  
       ‘[reveal] [one’s intentions]’

Although having a one-word interpretation for an idiom consisting of a head plus dependent(s) would always render it nondecomposable, having a multi-constituent interpretation does not automatically mean that the idiom is decomposable. In (5c), for example, despite the fact that the idiom’s interpretation consists of more than one constituent and despite its partial correspondence with the idiom’s constituent structure (i.e., *cards* corresponds to *intentions*), the structural isomorphism defined in (3) does not hold as [put] [on the table] is mapped to ‘reveal’, hence, the idiom is nondecomposable. It is important to emphasize that when diagnosing decomposability, it is crucial to use a literal, non-figurative paraphrase, rather than use another idiomatic/figurative one, as stated in (3). Figurative paraphrases are not reliable regarding the semantic composition of the idiom.<sup>1</sup>

Important for the present study are the results of an earlier experiment that examined the relevance of decomposability defined as in (3) for adults, by means of a completion task. The experiment revealed that adults’ performance on decomposable idioms was significantly better than their performance on nondecomposable idioms across both items and participants (Fadlon, Horvath & Siloni 2014). Decomposable idioms, unlike nondecomposable idioms, lend themselves to semantic composition of the figurative pieces. Fadlon, Horvath & Siloni (2014) argue that semantic composition facilitates the task of idiom retrieval.

As mentioned above, the dimension of decomposability has been argued to play a role in idioms’ level of flexibility (Nunberg, Sag & Wasow 1994; Rieheman 2001). However, flexibility does not seem to constitute a valid diagnostic for decomposability; it is merely a tendency, rather than a strict correlation, as shown in the literature regarding topicalization (Webelhuth & Ackerman, 1999), passivization (Bargmann & Sailer 2015), and verb second (Schenk 1995). Decomposability has also been argued to be relevant for idiom processing (Gibbs, Nayak & Cutting 1989) and acquisition (Gibbs 1987; 1991; Levorato & Cacciari 1999), but it has mostly not been carefully separated from “transparency” of idioms or from the “predictability” of their meaning, that is, from the extent to which the idiom’s meaning is transparent, or predictable based on its parts. Following Nunberg, Sag & Wasow’s original definition, we set aside transparency; in the present study, the meaning of all idioms, both decomposable and nondecomposable ones, cannot be predicted in advance based on the semantic composition of their subparts.

<sup>1</sup> An anonymous Glossa reviewer wonders whether the idiom ‘kick the bucket’ can also be paraphrased as ‘leave this life/stop living’ and consequently be classified as decomposable. The paraphrase ‘leave this life’, according to us, is not a literal paraphrase of (5a), but rather a metaphor comparing ‘dying’ to ‘leaving’; the paraphrase ‘stop living’ is not a straightforward paraphrase of the idiom itself, but rather a (somewhat artificial-sounding) paraphrase of the idiom’s direct paraphrase ‘die’. In a similar vein, one might raise doubts, suggesting that the idiom ‘put X’s cards on the table’ (5c) can be paraphrased by a (literal) multi-word expression, such as ‘put...in plain sight’, or ‘put...in the public domain’, and the idiom ‘take X with a grain of salt’ (4c) by a single word, such as ‘question’ or ‘distrust’. However, ‘put’ in the paraphrase of (5c) is employing it not in its literal meaning, but rather metaphorically, as if one physically puts his intentions somewhere, while in fact the combination ‘put...on the table’ means ‘reveal’, ‘expose’, when using a literal (nonmetaphoric) paraphrase. As for ‘take X with a grain of salt’ (4c), it means ‘accept something with reservation’ (because you consider it to be not completely true or right). Paraphrasing it as ‘distrust’, ‘question’, or ‘doubt’ misses the complex nuanced meaning of the idiom, which is not used to mean doubting something completely but rather to mean viewing/accepting something with reservation.

Given the centrality of the notion of decomposability in the idiom literature, and in light of its significance for idiom completion by adults, its potential impact on the acquisition of idioms by children is important to examine. To understand what the impact may be, we should first discuss (non)decomposability in relation to the storage of idioms.

## **1.2 Decomposability and storage**

Given the particular dual nature of idioms, i.e., their idiosyncratic meaning, which is associated with phrases involving syntactic structure, idioms need to be stored.

Nunberg, Sag & Wasow (1994) claim that the decomposability distinction corresponds to a difference in the way idioms are stored in the lexicon. Specifically, their proposal is that nondecomposable idioms are stored as phrasal constructions (whole VP collocations with their own idiosyncratic meaning), forming their own lexical entries. Decomposable idioms, in contrast, are assumed to get lexicalized after being analyzed into interpretationally interdependent words, which combine according to the general principles by which heads impose lexical (selectional/subcategorization) restrictions on their dependents.

In contrast to Nunberg, Sag & Wasow's (1994) split storage hypothesis, most of the theoretical literature, both earlier and subsequent proposals, assume uniform storage for phrasal idioms, considering decomposable as well as nondecomposable idioms as contextual restrictions on heads, that is, assuming both types of phrasal idioms to be stored based on their lexical formatives (see Everaert 2010 for an overview of this literature).

Empirical evidence supporting uniform storage of idioms based on L(exical)-selection by their head (see Everaert 2010) for both types of idioms (and against storage as independent phrasal entries) is provided by Horvath & Siloni (2009) and Fadlon, Horvath, Siloni & Wexler (to appear), who report surveys examining the cross-diathesis distribution of verb phrase idioms in Hebrew and English, respectively, as explained below.

Dubinsky & Simango (1996), Marantz (1997), and Ruwet (1991) report that in English, French & Chichewa there do not seem to be any idioms specific to the verbal (eventive) passive, while there are idioms specific to the adjectival (stative) passive. An idiom in the verbal passive must have a transitive version. Horvath & Siloni (2009) and Fadlon et al. (2016) confirm empirically the lack of both decomposable and nondecomposable phrasal idioms unique to the verbal passive, and propose an account for this robust generalization in terms of idiom storage, as follows.

It is common practice to assume that verbal (eventive) passives are formed in the syntax, i.e. beyond the storage component (Baker, Johnson & Roberts 1989; Collins 2005; Horvath & Siloni 2008; Meltzer-Asscher 2012, among others). It follows that there are no lexical entries that are passive verbs. If idioms must be stored under their head's lexical entry ("subentry storage"), then idioms specific to the verbal passive cannot be stored. We thus straightforwardly account for the fact that the verbal passive diathesis can only have idioms that are shared by the corresponding transitive (active) alternant. If nondecomposable verb phrase idioms were stored as independent lexical entries on their own, there would be no reason why nondecomposable idioms unique to the verbal passive should not exist; they would be stored as phrasal entries. The finding that neither decomposable, nor nondecomposable idioms can be unique to the verbal passive diathesis thus provides empirical evidence that both subtypes of idioms are listed in the lexicon as subentries of their head.

In sum, there are good reasons to assume that decomposable as well as nondecomposable verb phrase idioms are stored in a uniform way, as subentries, in the adult lexicon. This is not at odds with the results of the experiment mentioned above, which shows that adults retrieve decomposable idioms significantly better than nondecomposable ones. Although both types are stored the same way, decomposable idioms are retrievable



by semantic composition of the metaphoric pieces, which arguably facilitates retrieval. Nondecomposable idioms cannot have recourse to such retrieval. Storage as independent entries (multi-word units) is argued by Horvath & Siloni (2016) to exist only for idioms headed by a sentential functional head (“clausal” idioms) and idioms without recognizable internal syntactic structure.

Our predictions with regard to the effect of decomposability on children’s knowledge of idioms are given in the following two paragraphs.

If children, like adults, store all phrasal idioms as subentries, their performance would depend on whether or not they retrieve idioms more easily by semantic composition of the figurative pieces. If they do, then they would retrieve decomposable idioms significantly better than nondecomposable ones, on a par with adults. If the ability to use semantic composition is not developed enough to aid idiom retrieval, we would see no difference in performance between decomposable and nondecomposable idioms.

Recall that in nondecomposable idioms, as opposed to decomposable ones, the matching between the idiom’s constituents and the constituents of its literal paraphrase is either nonexistent or only partial (Section 1.1). It is possible that at certain stages of acquisition, this property hinders the ability to make use of the existing constituent structure of nondecomposable idioms for storage purposes. If so, then at these stages, children would not be able to store nondecomposable idioms as subentries. Rather, they would be compelled to store them in one piece, as independent entries. The question is: What consequences would this have regarding children’s performance? Recent literature (e.g., Tomasello 2003; Bannard & Matthews 2008; Arnon 2010; 2011) has argued that children in the early stages of acquisition attend to and store larger, multi-word units, rather than segment the linguistic input into its atoms (words, morphemes), and that these larger chunks have a facilitative effect in production. If this is indeed so, and if children store nondecomposable idioms as independent units, while decomposable ones have to be composed from the idiom’s subparts, children would show better performance with nondecomposable idioms as compared to decomposable ones (at least in production).

As mentioned in Section 1, previous findings have repeatedly revealed a discrepancy between idiom comprehension and production (Cacciari & Levorato 1989; Levorato & Cacciari 1992; 1995). In light of that, we examined comprehension and production separately, expecting our results to replicate this discrepancy.

The paper is structured as follows. In Sections 2 and 3 we present the two experiments we have conducted. The experiments were designed to examine whether and how the (non)decomposability of an idiom and the (non)existence of an open slot in it affect the performance of Hebrew-speaking children. Experiment 1 was dedicated to studying the effects of these variables on comprehension (section 2) and Experiment 2 was dedicated to studying their effects on production (section 3). The basic design of the experiments employed the designs of previous studies. (Gibbs 1987; Nippold & Tarrant Martin 1989; Nippold & Taylor, 1995; Levorato & Cacciari 1992; 1999; Levorato, Nesi & Cacciari 2004; among others); accordingly, as will be further detailed below, we used a multiple-choice task in the comprehension experiment and a completion task in the production experiment. Previous studies have also shown that children perform better if context is provided (Gibbs 1987; Cacciari & Levorato 1989; Levorato & Cacciari 1999; Laval 2003, among others). Therefore, in both tasks idioms were preceded by context in order to achieve performance at a level that would allow detecting potential differences in performance patterns between the four structural conditions. In section 3.4, we discuss the effects of the structural variables we manipulated and offer an analysis in terms of the developmental course of idiom acquisition. Our findings are summarized in section 4.

## 2 Experiment 1: Idiom comprehension in Hebrew

### 2.1 Participants

90 Hebrew speaking, monolingual children with no known linguistic or cognitive impairments took part in this experiment; 30 first graders (age range: 6–7, mean age: 6.23), 30 second graders (age range: 7–8, mean age: 7.67) and 30 third graders (age range: 8–9.5, mean age: 8.5). All participants were pupils of two public schools in central Tel-Aviv, constituting a homogenous population of upper middle class families. They were recruited and tested in their schools.

### 2.2 Method

#### 2.2.1 Stimuli

Stimuli consisted of 20 Hebrew verb phrase idioms composed of a transitive verb, an NP, and a PP. 5 idioms were full decomposable, 5 full nondecomposable, 5 open-slot decomposable and 5 open-slot nondecomposable. All idioms had a plausible literal meaning. Specifically, they did not involve selectional restriction violations. Thus, idioms such as *gava libo* ‘his heart got tall’ (idiomatic: ‘he became proud’) were excluded. Given the above restrictions on the type of idiom to be used and the need to counterbalance frequency, we had to include several idioms whose noun phrases involve a modifier. Idioms were classified with respect to decomposability by 6 Hebrew speaker graduate students of linguistics (age range 24–31), based on the definition of decomposability used here (3). These speakers were presented with the stimuli; they were asked to (i) provide a literal paraphrase and (ii) determine (based on (3) above) whether or not the constituent structure of the idiom (head and its dependents) corresponds to the constituent structure of its literal paraphrase. Idioms on which all speakers, including the two native speaker authors, agreed (regarding (i) and (ii)) were used as stimuli. Idioms such as *ataf et X be-cemer gefen* (‘wrapped X in-cotton.wool’), which was paraphrased either as ‘protected X to an exaggerated degree’ or as ‘spoiled X’ were not included. The open slot was in the pre-final XP position. We followed Levorato & Cacciari (1992; 1995; 1999) and used subjective frequency estimations provided by adults, in order to assess children’s exposure to the idioms and control for it between conditions. The measure of subjective frequency estimations is used when objective measures of frequency are lacking (Brysbaert & Cortese 2010), as is the case with Hebrew idioms. Subjective frequency estimations had been shown to reliably predict reaction times in idiom comprehension tasks (Libben & Titone 2008; Bonin, Méot & Bugajska 2013), and to correlate with objective frequency ratings (Nordmann & Jambazova 2017). 68 adult Hebrew speakers (age range: 19–44, mean age: 22) were asked to rate the frequency of 55 idiomatic phrases on a five point scale. Only phrases whose median ratings ranged between 3 and 5 were included in the research. Frequency was accordingly counterbalanced across the open-slot and the full idioms as well as between the decomposable and the nondecomposable ones. For the full list of idioms and their properties see Appendix A. Idioms did not include high register words and their meaning was not transparent (not close to literal meaning), based on judgments provided by 6 speakers (age range 22–29), in addition to the two native speaker authors. Thus, idioms judged too transparent like *avar al X be-štika* ‘passed over X in-silence’ (‘accepted something unpleasant without reacting/protest’) were excluded.<sup>2</sup>

<sup>2</sup> See Nippold and Taylor (2002) and Cain, Towse and Knight (2008) (among others) for the claim that the extent to which idioms’ figurative and literal meanings are close affects children’s performance.

### 2.2.2 Design

20 illustrated short stories providing the appropriate setting for an idiom without revealing its meaning were individually read out to participants. The idiom appeared at the end of the story, and was followed by a multiple-choice question. Stories were pseudo-randomly ordered to avoid a sequence of more than two idioms involving the same variables (i.e. full decomposable, full nondecomposable, open-slot decomposable, open-slot nondecomposable).

Subjects were asked to choose the correct interpretation among: (i) the correct idiomatic meaning; (ii) a literal, contextually inappropriate meaning; (iii) a contextually appropriate invented idiomatic meaning. To make sure participants were paying attention, they were presented with simple, multiple choice content questions once every 2–3 items. As a pilot session revealed a strong tendency of choosing the interpretation presented last, in the experimental session the answers were pseudo-randomly ordered not to include the correct idiomatic interpretation as the third option. Accordingly, 50% of the third options featured the literal interpretation and the other 50% the invented idiomatic interpretation. The correct idiomatic interpretation appeared either first or second. The contextually appropriate invented idiomatic option, which only knowledge of the idiom could rule out, were featured first in 25% of the items, second in 25% and last in 50%. Participants thus were not able to employ a heuristic strategy for responding.

The following is a translated example of the story and task constructed to test the comprehension of the idiom ‘to add oil to the fire’, the Hebrew equivalent of the English idiom ‘to add fuel to the fire’.

- (6) El’ad and his sister Shira sometimes argue about whose turn it is to use the computer. Their little sister Galit watches their quarrels and finds them amusing. Yesterday, for example, Shira was very angry with El’ad; she claimed that even though it was her turn to use the computer, El’ad didn’t let her take her turn until it was too late and they had to go to bed. El’ad said it wasn’t intentional and that he didn’t notice how late it was. Little Galit knew El’ad was lying because she saw him checking his watch and continuing to play his game. But she didn’t say anything since Mom told her: Galit, please don’t add oil to the fire.

In the story, what does it mean ‘to add oil to the fire’?

1. make the fight get worse (correct idiomatic)
2. pour oil on a flame (literal)
3. insult someone (invented idiomatic)

### 2.2.3 Procedure

Each subject participated in an individual session conducted by an instructor. The duration of each session was 20–30 minutes. The participant and the instructor sat side by side at a table. The instructor explained that she was about to read short stories and ask a question after each one. Once the child had expressed consent, the instructor placed a booklet containing the stories and tasks on the table. The instructor then read each story to the participant and presented the task orally. Once every 5 items, participants got a sticker they chose from a pack of assorted stickers. The group of first graders received a more valuable reward after the session was successfully completed: an animal shaped sketch board (instead of a sticker).

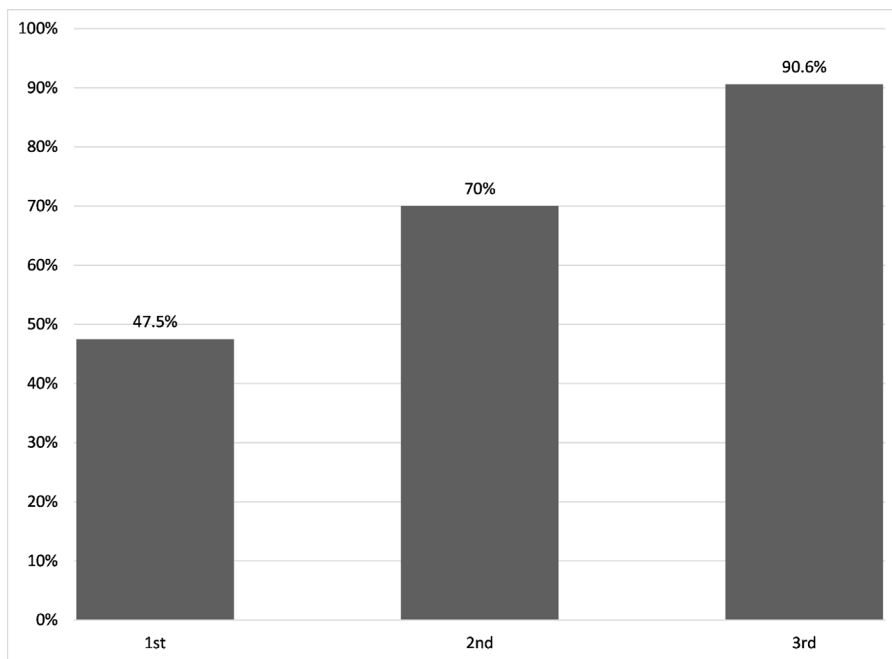
## 2.3 Results

Overall percentage of correct answers is shown in Figure 1, percentage of correct answers by decomposability and open-slot/full are shown in Figures 2 and 3, respectively. A by-participants  $3 \times 2 \times 2$  repeated measures ANOVA with the within-subject factors “decompos-

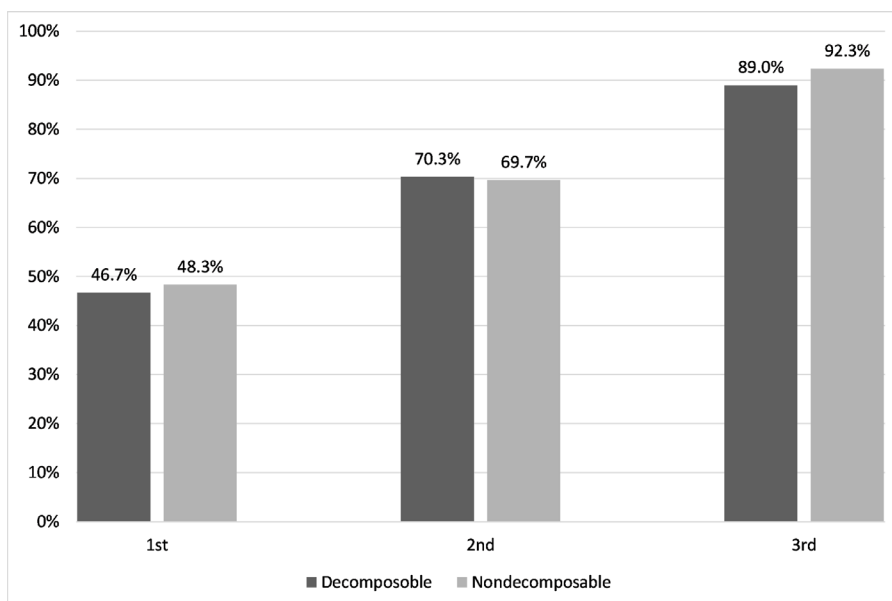


ability” (decomposable/nondecomposable) and “existence of an open slot” (open-slot/full) and the between-subjects factor “age group” (1<sup>st</sup> grade/2<sup>nd</sup> grade/3<sup>rd</sup> grade) was carried out. This analysis yielded a significant effect for age ( $F(2, 87) = 45.7, p < .001$ ). A by-item analysis with the within-item factor ‘age group’ and the between-item factors “decomposability” and “existence of an open slot” also revealed a significant main effect for age ( $F(2, 32) = 112.05, p < .001$ ). No other main effects or interactions were found ( $ps > .27$ ).

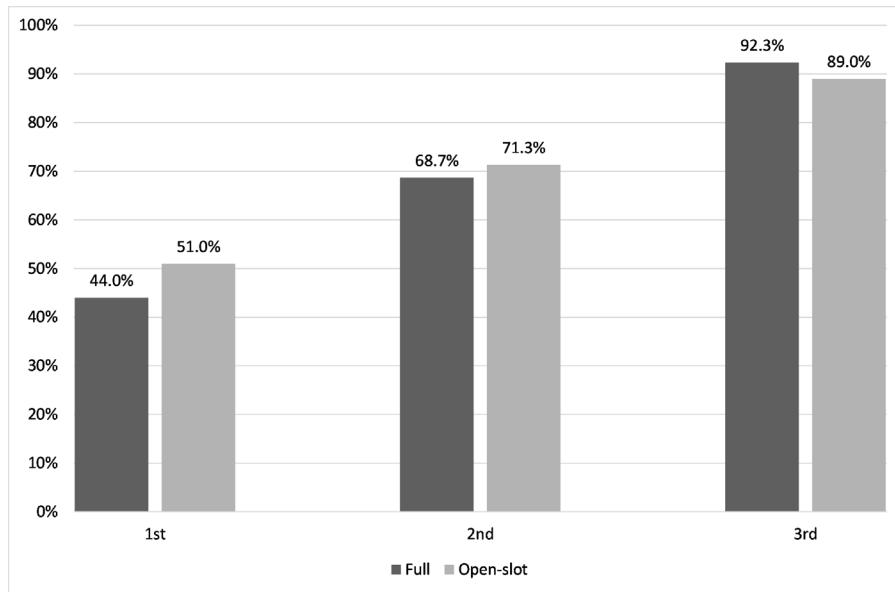
Post hoc analyses revealed the effect for age to be monotonic: second graders ( $M = 14, SD = 3.78$ ) performed significantly better than first graders ( $M = 9.5, SD = 4.6$ ) (t-test for correlated samples comparing average score per idiom:  $t(19) = 9.6, p < .001$ ) and third graders ( $M = 18.1, SD = 0.89$ ) performed significantly better than second graders (t-test for correlated samples comparing average score per idiom:  $t(19) = 4.18, p < .001$ ).



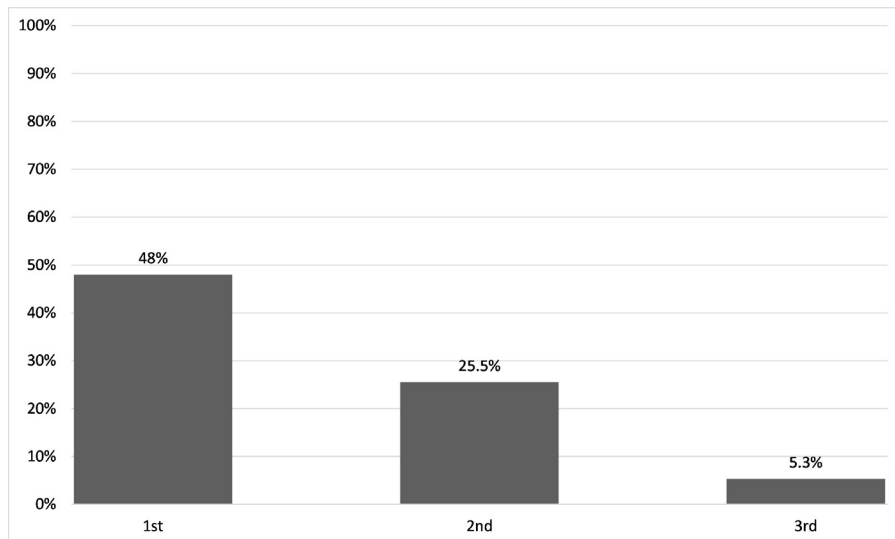
**Figure 1:** Percentage of correct choices per grade.



**Figure 2:** Percentage of correct choices per grade by decomposability.



**Figure 3:** Percentage of correct choices per grade by full/open-slot.

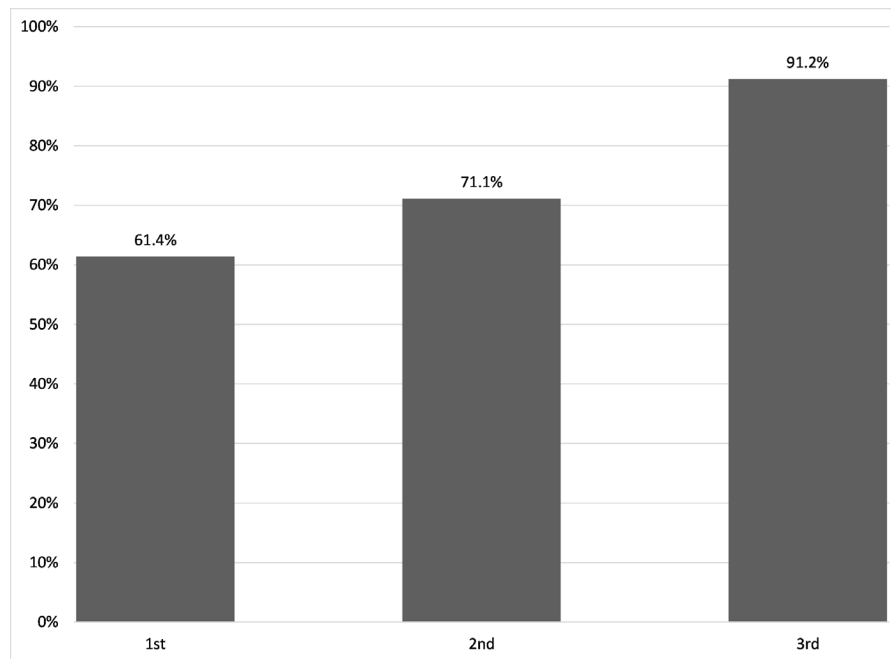


**Figure 4:** Percentage of literal errors out of all errors per grade.

We tested each group to see if the results reflected an above chance level, which would be 1/3 given the 3 possible responses. It turns out that even the first graders showed above-chance performance: First grade: 47.5%, single sample t-test:  $t(29) = 3.4, p = .001$ , second grade: 70%, single sample t-test:  $t(29) = 10.46, p < .001$ , and third grade: 90.6%, single sample t-test:  $t(29) = 70.3, p < .001$ .

Part of the improved performance over age is due to the decrease in literal errors. 74.8% of the responses of first graders were one of the 2 idiomatic interpretations, as were 91.5% of the responses of second graders and 99.5% of the responses of third graders. Figure 4 shows that the percentage of literal errors out of all errors (the alternative error is the choice of a wrong idiom interpretation) is almost half in the first grade, but has decreased to 5.3% by third grade.

It makes sense to attribute the literal errors to the lack of knowledge of (or inability to process) figurative language. The literal readings made no sense in context; we assume that if a child could use figurative language to analyze a particular sentence, she would have chosen one of the idiomatic interpretations, either pick the correct one (if she knows the idiom) or make a random choice between the correct and invented idiom (if she doesn't know the idiom).



**Figure 5:** Percentage of correct choices out of the two idiomatic responses.

However, we cannot attribute all the development in performance to the ability to use figurative language. Figure 5 shows that if a child has chosen one of the 2 idiomatic interpretations (correct or invented), the older the child, the more likely that this choice is the correct idiom. The choice of the correct idiom, given that an idiomatic meaning has been chosen, increases with each age group. 91.2% of the third graders' choice of idiomatic interpretations (out of the 2 possibilities) are correct, whereas only 61.4% of the first graders' choice of idiomatic interpretations are correct.

Given that a child has decided that the literal meaning makes no sense and that a figurative meaning is needed, the chance level of responding, assuming that a child has no knowledge of the idiom itself, is  $\frac{1}{2}$ . We tested the percentage of choice of the correct idiom for each age group against this chance level of .5 using a single sample one-tailed t-test. In each case, even for the first graders, the results showed that children knew the correct idiom at an above chance level. First grade: 61.4%, single sample t-test:  $t(29) = 3.34$ ,  $p = .001$ , second grade: 71.1%, single sample t-test:  $t(29) = 8.99$ ,  $p < .001$ , and third grade: 91.2% single sample t-test:  $t(29) = 55.94$ ,  $p < .001$ .

#### **2.4 Interim discussion**

As demonstrated above, two simultaneous developmental processes were observed: first, there is a gradual decrease in the rate of literal errors. This is consistent with (i) findings regarding Italian, French and English (Ackerman 1982; Levorato & Cacciari 1999; Laval 2003; Levorato, Nesi & Cacciari 2004), (ii) studies showing that children's knowledge of metaphors and similes matures gradually over early school years in tandem with the decrease in the tendency to attribute literal interpretations to figurative phrases (Vosniadou 1987; Winner 1988).

Second, there is a development reflecting knowledge of particular idioms. Just as in grammar, a child has to develop both an underlying grammatical ability (e.g. syntax, semantics) and the knowledge of particular lexical items. Interestingly, we have observed in experiment 1 both of these processes in action. Both the general ability to use figurative language and knowledge of particular idioms increase in the age-range that we have studied. Our decision to use a 3-choice experiment, in which the child can make an error either by choosing a literal rather than figurative interpretation or by choosing the wrong

figurative (wrong idiomatic) meaning, allowed for us to study both of these processes in the same experiment.

Existence of an open slot and decomposability seem irrelevant to idiom comprehension in the early school years: Neither had an effect on children's performance.

### 3 Experiment 2: Production

#### 3.1 Participants

Participants were 90 Hebrew speaking, monolingual children (from the same two public schools as for experiment 1), with no diagnosed cognitive or linguistic deficits, who did not take part in experiment 1: 30 first graders (age range: 6–7, mean age: 6.41), 30 second graders (age range: 7–8, mean age: 7.64) and 30 third graders (age range: 8–9.5, mean age: 8.61).

#### 3.2 Method

##### 3.2.1 Stimuli

Stimuli consisted of the same list of idioms used in experiment 1.

##### 3.2.2 Design

20 short illustrated stories providing the appropriate setting for the use of each idiomatic phrase were composed. The idiom appeared at the end of the story in an incomplete form that allowed the recognition of the idiom but not a correct guess; in each idiom one content word was omitted. The position of the omitted element was determined based on the following criteria: The element was (i) a lexical head (not a functional one, which could be guessed); (ii) not a constituent more easily completed based on context, e.g. not *table* in *put one's cards on the table*; (iii) not a word that occurred in more than one idiom (e.g., *katan/ktana* 'small.ms/fm'). The position of the missing element was either the first or second constituent to match these criteria.

Subjects were asked to complete the idioms. As in experiment 1, participants were presented with a content question structurally similar to the target task (i.e. a fill-in-the-blank task) once every 2–3 items, to make sure they were paying attention, and items were pseudo-randomly ordered to avoid a sequence of more than two idioms involving the same variables. However, unlike in experiment 1, stories did provide the meanings of the idioms in order to facilitate the retrieval of the target idiom.

The following is a translated example of the story and task constructed to elicit the production of the idiom 'add oil to the fire'. Notice that apart from the lines which provide the meaning of the target idiom (appearing here in boldface), the story is similar to the one used in experiment 1 (see (6) above).

- (7) El'ad and his sister Shira sometimes argue about whose turn it is to use the computer. Their little sister Galit watches their quarrels and finds them amusing. Yesterday, for example, Shira was very angry with El'ad; she claimed that even though it was her turn to use the computer, El'ad didn't let her take her turn until it was too late and they had to go to bed. El'ad said it wasn't intentional and that he didn't notice how late it was. Little Galit knew El'ad was lying because she saw him checking his watch and continuing to play his game. **Galit wanted to interfere and say "Shira is right! I saw El'ad checking his watch" but she didn't get a chance since Mom asked her to keep quiet and not add oil to the \_\_\_\_.**

##### 3.2.3 Procedure

The procedure was identical to the one of experiment 1. Each subject participated in an individual session conducted by an instructor. The duration of each session was 20–30

minutes. The participant and the instructor sat side by side at a table. The instructor explained that she was about to read short stories which end with an incomplete sentence that requires completion. Once the child had expressed consent, the instructor placed a booklet containing the stories on the table. The instructor then read each story to the participant, uttering an *m* sound to indicate the blank part. Once every 5 items, participants got a sticker they chose from a pack of assorted stickers. The group of first graders received a more valuable reward after the session was successfully completed: an animal shaped sketch board (instead of a sticker).

Each response that matched the target word was given the score of 1. The rest of the responses were viewed as incorrect. More details about the types of errors we observed are given in Appendix B.

### 3.3 Results

The accuracy of participants' responses was analyzed using a by-participants  $3 \times 2 \times 2$  repeated measures ANOVA with the within-subject factors "decomposability" (decomposable/non-decomposable) and "existence of an open slot" (open-slot/full) and the between-subjects factor 'age group' (1<sup>st</sup> grade/2<sup>nd</sup> grade/3<sup>rd</sup> grade). This analysis yielded a significant effect for age group ( $F(2, 87) = 45.53, p < .001$ ), decomposability ( $F(1, 87) = 98.08, p < .001$ ) and existence of open slot ( $F(1, 87) = 8.926, p = .004$ ), as well as significant interactions between decomposability and existence of an open slot ( $F(1, 87) = 38.84, p < .001$ ) and between these two variables and age group ( $F(2, 87) = 10.06, p < .001$ ). A by-item analysis with the within-item factor "age group" and the between-item factors "decomposability" and "existence of an open slot" revealed a significant main effect for age ( $F(2, 32) = 36.5, p < .001$ ) and decomposability ( $F(1, 16) = 9.298, p = .008$ ). Average success rates and standard deviations by grade and manipulated factors are provided in Table 1.

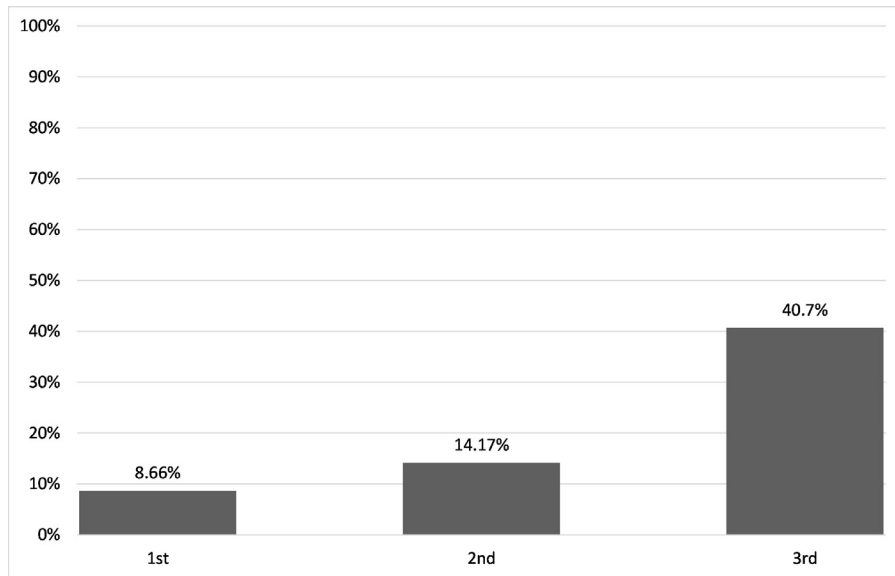
As shown in Figure 6, planned comparisons revealed a monotonic effect for age, second graders ( $M = 4.25$  (14.17%),  $SD = 5.3$ ) performed significantly better than first graders ( $M = 2.6$  (8.66%),  $SD = 3.33$ ) (t-test for correlated samples comparing average score per idiom:  $t(19) = 2.71, p = .014$ ) and third graders ( $M = 12.21$  (40.7%),  $SD = 7.2$ ) performed significantly better than second graders (t-test for correlated samples comparing average score per idiom:  $t(19) = 4.92, p < .001$ ).

Planned comparisons of performance with decomposable idioms vs. nondecomposable idioms revealed that for all age groups performance with nondecomposable idioms was significantly better than the performance with decomposable idioms. For the group of first graders the average success rate with nondecomposable idioms was 15.6% vs. 1.6% with decomposable idioms (t-test for correlated samples:  $t(29) = 4.89, p < .001$ ), the group of second graders averaged 25% with nondecomposable idioms and 3.33% with decomposable idioms (t-test for correlated samples:  $t(29) = 7.13, p < .001$ ) and for the group of third graders the average success rate with nondecomposable idioms was 50.3% vs. 31% with decomposable idioms (t-test for correlated samples:  $t(29) = 5.6, p < .001$ ).

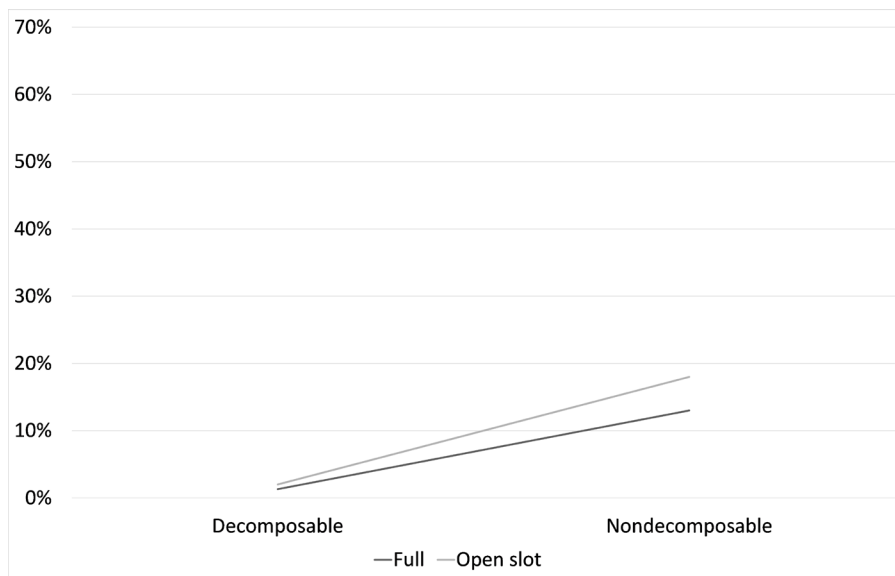
**Table 1:** Average success and standard deviation per grade by full/open-slot and decomposability.

Idioms		1 <sup>st</sup> grade		2 <sup>nd</sup> grade		3 <sup>rd</sup> grade	
		Mean	SD	Mean	SD	Mean	SD
full	Decomposable	1.3%	.03	3.3%	.51	38%	.24
	nondecomposable	13%	.13	18%	.19	39.3%	.22
open-slot	Decomposable	2%	.03	3.3%	.51	26%	.20
	nondecomposable	18%	.09	32%	.17	61.3%	.15





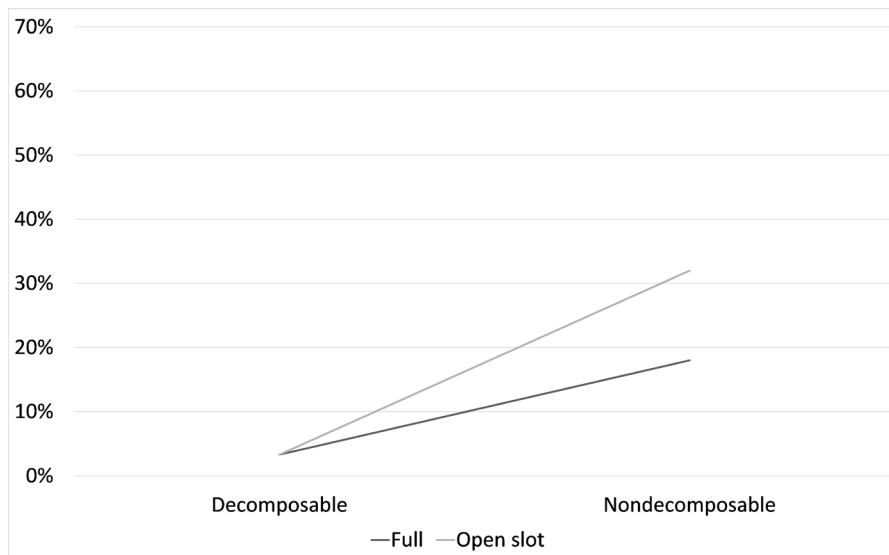
**Figure 6:** Percentage of correct completions per grade.



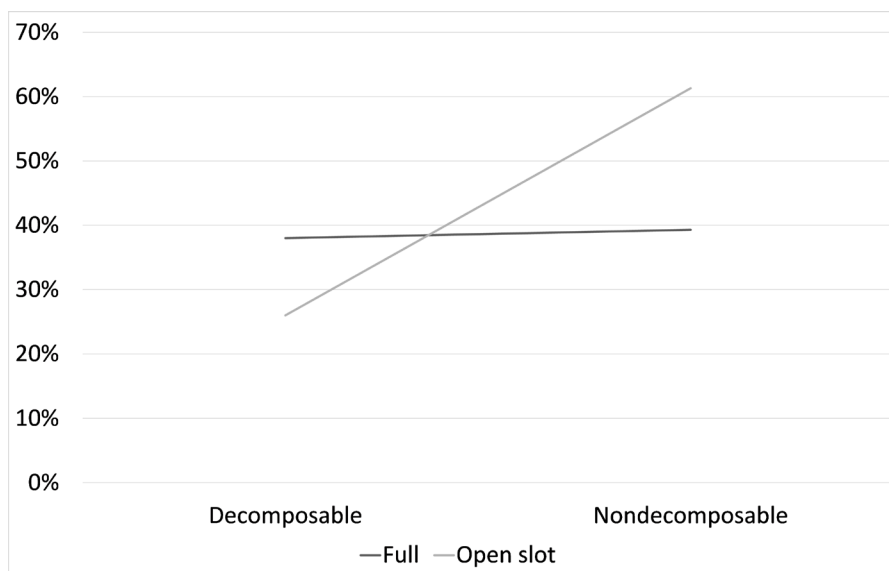
**Figure 7:** Decomposability and existence of an open slot – first grade.

With regard to existence of an open slot, however, planned comparisons revealed that the difference between open-slot and full idioms only affected the performance of second graders. As shown in the next paragraph, this result is due to different performance being observed only with nondecomposable (open-slot/full) idioms. For the group of first graders the average success rate with open-slot idioms was 10% vs. 7.1% with full idioms (t-test for correlated samples:  $t(29) = 1.01, p = .31$ ), the group of second graders averaged at 17% with open-slot idioms and 10% with full idioms (t-test for correlated samples:  $t(29) = 2.9, p = .006$ ), and for the group of third graders the average success rate with open-slot idioms was 43% vs. 37.6% with full idioms (t-test for correlated samples:  $t(29) = 1.81, p = .08$ ).

Post hoc analyses focused on revealing the source of the interaction between decomposability and existence of an open slot found a significant interaction between these variables only for the group of second graders ( $F(1, 29) = 9.2, p = .005$ ) and the group of third graders ( $F(1, 29) = 34.776, p < .001$ ), but not for the group of first graders ( $F(1, 29) = 1.64, p = .31$ ). As Figures 7 and 8 show, the existence of an open slot improved



**Figure 8:** Decomposability and existence of an open slot – second grade.



**Figure 9:** Decomposability and existence of an open slot – third grade.

second graders' chances of retrieving the correct response for nondecomposable target idioms (18% success rate with full vs. 32% success rate with open-slot, t-test for correlated samples:  $t(29) = 3.252$ ,  $p = .003$ ), but, as both full decomposable idioms and open-slot decomposable ones averaged 3.3%, it did not affect their performance with decomposable idioms at all.

As far as third graders are concerned, the interaction was disordinal (Figure 9): like with second graders, existence of an open slot improved their chances of retrieving the correct response for nondecomposable idioms. However, lack of an open slot (full idioms) improved their chances of producing the correct response for decomposable idioms. Post-hoc analyses separately comparing third graders' performance with full and open-slot idioms within two levels of decomposability revealed this pattern to be significant. Third graders' average success rate with full decomposable idioms was 38% vs. 26% with open-slot decomposable idioms (t-test for correlated samples:  $t(29) = 2.82$ ,  $p = .008$ ), while their average success rate with full non-decomposable idioms was 39.3% vs. 61.3% with open-slot nondecomposable idioms (t-test for correlated samples:  $t(29) = 4.74$ ,  $p < .001$ ).

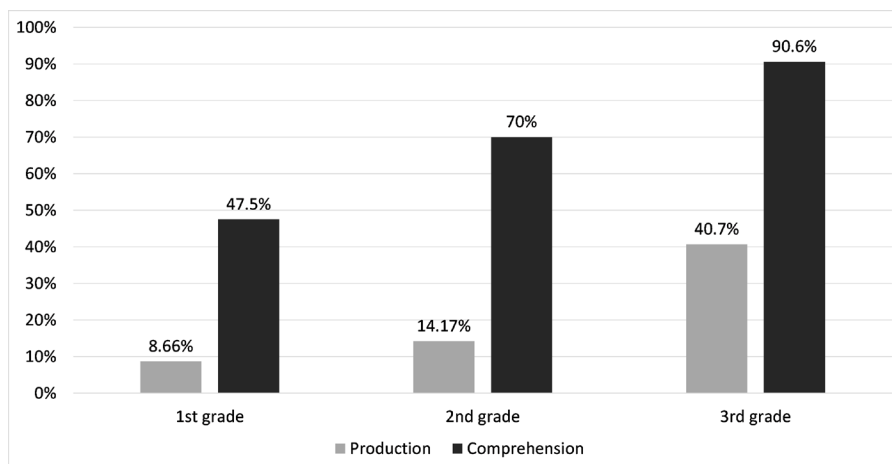
Additional analyses, separately comparing the three groups' performance with non-decomposable vs. decomposable idioms within the set of full idioms and within the set of open-slot idioms, revealed that within the set of full idioms performance with nondecomposable idioms was significantly better only for 1<sup>st</sup> and 2<sup>nd</sup> graders. The average success rate with full nondecomposable idioms was 13% vs. 1.3% with full decomposable idioms for 1<sup>st</sup> graders (t-test for correlated samples:  $t(29) = 3.52, p = .001$ ). For 2<sup>nd</sup> graders, average success rate with full nondecomposable idioms was 18% vs. 3.3% with full decomposable idioms (t-test for correlated samples:  $t(29) = 4.85, p < .001$ ). In contrast, third graders' average success rate with full nondecomposable idioms was 39.3% vs. 38% with full decomposable idioms (t-test for correlated samples:  $t(29) = 0.4, p = .7$ ).

Within the set of open-slot idioms, on the other hand, performance with nondecomposable idioms was significantly better for all three age groups. For 1<sup>st</sup> graders, average success rate with open-slot nondecomposable idioms was 18% vs. 2% with open-slot decomposable idioms (t-test for correlated samples:  $t(29) = 4.55, p < .001$ ). For 2<sup>nd</sup> graders, average success rate with open-slot nondecomposable idioms was 32% vs. 3.3% with open-slot decomposable idioms (t-test for correlated samples:  $t(29) = 64.1, p < .001$ ). For 3<sup>rd</sup> graders, average success rate with open-slot nondecomposable idioms was 61.3% vs. 26% with open-slot decomposable idioms (t-test for correlated samples:  $t(29) = 6.41, p < .001$ ).

### 3.4 Discussion

As in experiment 1 (comprehension), the results of experiment 2 reveal a gradual maturation in the ability of Hebrew-acquiring children to produce idioms. However, as Figure 10 demonstrates, the average success rates observed in the comprehension experiment were considerably higher than the ones observed in the production experiment.

Given that discrepancies between what children seem to understand and their performance in tasks designed to elicit production are often observed in acquisition studies (see Clark & Hecht 1983 for an overview), including those targeting the acquisition of idioms (Cacciari & Levorato 1989; Levorato & Cacciari 1995), the divergence between idiom comprehension and production is not unexpected, as mentioned in Section 1.2. Retrieving the meaning of an idiomatic sequence of words aided by context is easier for acquirers than retrieving the idiomatic sequence itself based on meaning inferred from context, even when most items composing the idiom are given. Moreover, comprehension of idioms being easier than production, we do not expect the effects observed in experiment 2 to be necessarily attested in experiment 1.



**Figure 10:** Percentage of correct responses in production vs. comprehension per grade.

**Table 2:** Average success per grade by full/open-slot and decomposability.

Idioms		1 <sup>st</sup> grade	2 <sup>nd</sup> grade	3 <sup>rd</sup> grade
full	decomposable	1.3%	3.3%	38%
	nondecomposable	13%	18%	39.3%
open-slot	decomposable	2%	3.3%	26%
	nondecomposable	18%	32%	61.3%

In some of our items, labeled AddWord in Appendix A, the omitted word was part of a noun phrase which includes an additional lexical item, while in others, if we ignore the definite article, the omitted word constitutes a noun phrase on its own. Compare, for example, the AddWord item ‘put a healthy head in [*a sick bed*]’ to the non-AddWord item ‘buried the head in [*the sand*]’, (omitted words are italicized; additional lexical item in bold). In order to complete the first idiom, the child’s task is to retrieve the adjective *sick* which modifies the word *bed*, which is not omitted. Whereas in the case of the second idiom, the child’s task is to retrieve the word *sand* which projects the target noun phrase *the sand* without the involvement of additional material. This variability could, in principle, lead to higher success rates with AddWord idioms compared to non-AddWord idioms, since the former contains additional cues (‘bed’ in the above AddWord idiom) that could potentially facilitate the retrieval of the omitted part. However, further by-subjects analyses comparing success rates between AddWord and non-AddWord items for each age group demonstrate that this was not the case.

Along the decomposability vs. nondecomposability dimension and the open-slot vs. full idiom dimension, the specific interactions revealed in experiment 2 are summarized in (8) and (9) below, and discussed directly. The course of development is presented in Table 1 repeated as Table 2 (standard deviations omitted).

- (8) Decomposability vs. nondecomposability
- For all age groups, performance with nondecomposable idioms was significantly better than performance with decomposable idioms.
  - For third graders exclusively, within the set of full idioms, performance with nondecomposable idioms is not significantly better than with decomposable ones (39.3% vs. 38%, respectively).
- (9) Full vs. Open-slot
- For both second and third graders, within the set of nondecomposable idioms, performance with open-slot idioms was significantly better than with full ones (61.3% vs. 39.3%, respectively).
  - For third graders exclusively, within the set of decomposable idioms, performance with full idioms was significantly better than performance with open-slot idioms (38% vs. 26%, respectively).

### 3.4.1 Decomposability vs. Nondecomposability

(Non)decomposability played a role in production across age groups: Significantly superior scores for nondecomposable idioms in the production (completion) task were found across all age groups (8a). Recall that adults, in contrast, perform significantly better on decomposable idioms than on nondecomposable ones in a parallel retrieval (completion) task (as mentioned in Section 1.1). Notwithstanding, there is evidence that adults store decomposable and nondecomposable idioms the same way, by subentry storage, as discussed in Section 1.2 and illustrated in Table 3.

**Table 3:** Decomposable and nondecomposable idioms – adults.

Decomposable	Nondecomposable
[spill] [the beans] [divulge] [the secret]	shoot the breeze [chat]
<b>Subentry</b> storage (SPILL) SPILL <Ag> <Th> ‘cause/let flow’	<b>Subentry</b> storage (SHOOT) SHOOT <Ag> <Th> ‘eject by a sudden release’
SPILL’ <Ag> <Th → beans: <b>N<sub>PL+DEF</sub></b> > ‘divulge the secret’	SHOOT’ <Ag> <Th → breeze: <b>N<sub>SG+DEF</sub></b> > ‘chat’

**Table 4:** Decomposable vs. nondecomposable idioms – children.

Decomposable	Nondecomposable
[spill] [the beans] [divulge] [the secret]	shoot the breeze [chat]
<b>Subentry</b> storage (SPILL) SPILL <Ag> <Th> ‘cause/let flow’	<b>Independent</b> storage ‘shoot the breeze’ ‘chat’
SPILL’ <Ag> <Th → beans: <b>N<sub>PL+DEF</sub></b> > ‘divulge the secret’	

Given the evidence suggesting that adults store both decomposable and nondecomposable idioms by subentry storage, the advantage they show with decomposable idioms has to be attributed, as noted in Section 1.2, to the fact that decomposable idioms lend themselves to semantic composition of the figurative pieces, unlike nondecomposable idioms. In other words, adult performance indicates that semantic composition may facilitate the task of putting the idiomatic pieces together during idiom retrieval.

According to our prediction in Section 1.2, if first to third graders stored decomposable as well as nondecomposable idioms by subentry storage on a par with adults, it would be unexpected that they complete nondecomposable idioms more easily than decomposable ones. In contrast, if children at this stage store nondecomposable idioms as independent units (unlike decomposable idioms), as illustrated in Table 4, then the previously observed facilitative effect of larger information chunks on processing (Tomasello 2003; Bannard & Matthews 2008; Arnon 2010; 2011) mentioned in Section 1.2 predicts that they would be better at completing nondecomposable idioms (in comparison to decomposable ones). The observed performance pattern is therefore consistent with the split storage account, according to which children store nondecomposable idioms as independent units whereas decomposable ones are stored as subentries.

We suggest that the reason why children at this stage store nondecomposable, but not decomposable, idioms as independent entries (unlike adults) is their inability to reconcile the constituent structure of nondecomposable idioms with their lack of semantic composition. That is, in the absence of matching between the idiom’s constituents and the constituents of its literal paraphrase, children are unable to make use of the existing constituent structure of these phrases for storage, and therefore do not store these idioms as subentries. For instance, storing the idiom *shoot the breeze* as a subentry of *shoot* (see Table 3) may be challenging because there is no bit of the idiom’s meaning corresponding to the head *shoot*. We are not suggesting that children are unable to recognize that nondecomposable idioms have internal constituent structure just like literal phrases do. Rather, what we propose is that the absence of semantic composition matching the



idiom's constituent structure results at this stage of development in storage of nondecomposable idioms as independent (single unit) entries. Decomposable idioms, in contrast, involve no such mismatch of semantic composition and constituent structure, and thus, analyzed into their constituents, they get stored as subentries of the lexical entry of their head from the beginning.

In sum, we propose that the pattern found in children's production of decomposable vs. nondecomposable idioms follows from:

- (10) a. At these stages of acquisition, nondecomposable phrasal idioms are stored as independent units, while decomposable ones are stored as subentries of their lexical head, as illustrated in Table 4.
- b. Retrieval of independently stored entries is easier for children than retrieval of subentries, as the latter requires composition of the idiom from its subparts, which is more difficult at this stage.

Further, experiment 2 indicates that in third grade there is a change in the pattern of performance regarding full idioms: Performance on nondecomposable idioms is no longer significantly better than performance on decomposable ones. This results from a drastic improvement from second to third grade in performance with decomposable idioms, as can be seen in Table 2, and as is explained below.

Performance on both open-slot and full decomposable idioms improves considerably (3.3% to 26%, and 3.3% to 38%, respectively). But, clearly, the increase is bigger with full than with open-slot decomposable idioms. Nondecomposable idioms, in contrast, keep a relatively steady course of development, as is clear from the Table. Further, performance on open-slot nondecomposable idioms (61.3%) is significantly better than performance on all other types. Hence, despite the improvement that open-slot decomposable idioms show from second to third grade, in third grade, they still do not approach performance on open-slot nondecomposable ones. The discrepancies in performance on full vs. open-slot idioms are discussed in Section 3.4.2. We first explain the general improvement in performance with decomposable idioms.

We attribute this improvement to the maturation of third graders' ability to use semantic composition, as summarized in (11), and explained below.

- (11) The ability to retrieve (produce) idioms by semantic composition matures at third grade.

For adults, semantic composition is the standard strategy. Hence, as mentioned above, they are better at producing decomposable than nondecomposable idioms. At the beginning, children have a hard time producing idioms by semantic composition. They are much better at retrieval of one-piece stored units, as explained above. In third grade their ability to use semantic composition for retrieval of idioms improves considerably, as shown by the increase in percentage of correct responses on decomposable idioms.

In the course of development of idiom knowledge in children, the transition to the adult state is expected to involve in addition to the maturation process in (11), also the maturation of an additional ability. Recall we suggested that the reason why children at this stage store nondecomposable idioms as independent entries (unlike adults) is their inability to reconcile the constituent structure of nondecomposable idioms with their lack of corresponding semantic composition. This ability should mature. Its maturation in children's developing grammar will permit them to utilize constituent structure for lexical storage also in the absence of corresponding semantic composition. They will then

store nondecomposable idioms as subentries and exhibit change of performance in favor of decomposable idioms vs. nondecomposable idioms on a par with adults.<sup>3</sup>

In the next section we discuss the discrepancies in performance on full vs. open-slot idioms.

### 3.4.2 Full vs. Open-slot

Turning to the division between full vs. open-slot idioms, the question raised by our findings is: Why is there a difference between performance on full vs. open-slot idioms? Specifically, the following two questions arise from our findings.

- (12) a. Regarding the set of decomposable idioms, why does performance by third graders become significantly better on full idioms than on open-slot ones (38% vs. 26%, as shown in Table 2 and stated in (9b))?
- b. Regarding the set of nondecomposable idioms, why is performance on open-slot idioms significantly better than on full idioms, both for second graders (18% vs. 32%) and for third graders (39.3% vs. 61.3%), (see Table 2 and (9a))?

Starting with question (12a), the maturation of the ability to retrieve idioms by semantic composition (11) brings about considerable improvement in retrieval of decomposable idioms, both full and open-slot ones. However, the improvement with full idioms is more pronounced. We suggest the reason for this is the following: The existence of a mix of lexically fixed elements and open slots makes the stored subentries less uniform for open-slot idioms than for full ones. We attribute the difference in performance between full vs. open-slot decomposable idioms to this lack of uniformity in the subentry of open-slot idioms, which renders them harder to retrieve by semantic composition.

Question (12b) involves the existence of a reverse pattern: Within the set of nondecomposable idioms (in contrast to decomposable ones), it is the performance on open-slot idioms that is significantly better; open-slot idioms score significantly better than full ones, both for second graders (18% vs. 32%) and for third graders (39.3% vs. 61.3%), (see Table 2). Why would nondecomposable idioms exhibit the opposite pattern of asymmetry between performance on full vs. open-slot idioms?

If, as we propose, nondecomposable idioms are stored at these stages of acquisition by independent storage, no composition is involved in their retrieval. While for retrieval by composition, an open slot makes the task harder, due to the nonuniform subentry representation of the idiom (as just explained), for retrieval of a whole unit (an independent entry), the existence of an open slot makes the task apparently easier.

We suggest that the reason for this is that although the open-slot and the full idioms used in the experiment have the same number of constituents (three each), the open-slot idioms uniformly have fewer constituents with fixed lexical material than do full ones. The open-slot idioms consist of a verb plus one lexically fixed phrase (in addition to the open slot), whereas the full idioms consist of a verb plus two lexically fixed phrases. This suggests that in the case of independent storage, the more constituents with fixed lexical material an idiom has, the more difficult it is for children to retrieve it. This then seems to account for why second and third graders' performance is better on open-slot than on

<sup>3</sup> The maturation of this additional ability results in a shift in the storage of nondecomposable idioms from independent to subentry storage due to the fact that independent storage is costlier to the grammar than subentry storage. This is shown in Horvath & Siloni (2016); independent storage in the adult grammar is used only as a last resort, namely only for idioms that have no lexical head (lack internal structure or are headed by a functional category), and thus have no subentry storage option.

full nondecomposable idioms. Finally, first graders do not show this pattern of distinction: Performance on open-slot nondecomposable idioms is not significantly better than on their full counterparts (13% vs. 18%, respectively). This is so as their performance is too low to detect a significant difference.

#### 4 Conclusion

The study investigates the effects of decomposability and the existence of an open slot on the acquisition of verb phrase idioms by school-children (first to third graders) acquiring Hebrew. Neither decomposability nor the existence of an open slot influenced idiom comprehension. But idiom production was affected by both factors. Performance with nondecomposable idioms was significantly better than performance with decomposable idioms across age groups, in contrast with performance by adults, who score better on decomposable idioms. In third grade, however, the pattern changes: performance on nondecomposable idioms is no longer significantly better than on decomposable ones for idioms that are full. Further distinctions are observed for full vs. open-slot idioms: (i) As far as nondecomposable idioms are concerned, performance on open-slot idioms was significantly better than on full ones for both second and third graders. (ii) As far as decomposable idioms are concerned, an inverse pattern is observed: Performance on full idioms was significantly better than performance on open-slot idioms.

We propose that the effect of the decomposability factor results from the different storage of nondecomposable idioms vs. decomposable ones at this stage: While decomposable idioms are stored as subentries of their head (as in the adult grammar), nondecomposable ones are stored as independent units, unlike in the adult grammar. Since children retrieve stored units more easily than phrases that have to be composed, performance on nondecomposable idioms is better. Further, we propose that the reason why children at this stage store nondecomposable idioms by independent storage (unlike adults) is that they have difficulty reconciling the constituent structure of nondecomposable idioms with their lack of semantic composition. Finally, the interactions of decomposability and the existence of an open slot for second and third graders follow from the different modes of retrieval each storage strategy imposes: While for retrieval of subentries (by semantic composition), an open slot makes the task harder, due to the nonuniform lexical representation of the idiom, for independent unit storage, the more constituents with fixed lexical material an idiom has, the more difficult it is for children to retrieve it.

#### Abbreviations

ACC = accusative, Ag = agent, DEF = definite, PL = plural, SG = singular, Th = theme, X = open slot

#### Additional Files

The additional files for this article can be found as follows:

- **Appendix A.** Idioms used in experiments 1 and 2. DOI: <https://doi.org/10.5334/gjgl.542.s1>
- **Appendix B.** Error analysis, experiment 2. DOI: <https://doi.org/10.5334/gjgl.542.s1>

#### Acknowledgements

We would like to thank Jenny Birger, Dganit Kim and Nathaniel Sharon for their help with subject recruitment and data collection. We are also grateful to the audiences at the TAU-GU workshop on Relative Clauses and Idioms, Goethe University, Frankfurt, November, 28–30, 2014, and the 8th BCGL The Grammar of Idioms, CRISSP (KU Leuven), Brussels,

June 4–5, 2015, where earlier versions of this work have been presented. Finally, a very special thank you goes to the 180 patient and cooperative children who participated in this study. This research was supported by Grant No 2009269 from the United States–Israel Binational Science Foundation (BSF).

### Competing Interests

The authors have no competing interests to declare.

### References

- Ackerman, Brian P. 1982. On comprehending idioms: Do children get the picture? *Journal of Experimental Child Psychology* 33(3). 439–454. DOI: [https://doi.org/10.1016/0022-0965\(82\)90058-3](https://doi.org/10.1016/0022-0965(82)90058-3)
- Arnon, Inbal. 2010. *Starting big – The role of multi-word phrases in language learning and use*. Stanford, CA: Stanford University dissertation.
- Arnon, Inbal. 2011. Units of learning in language acquisition. In Inbal Arnon & Eve V. Clark (eds.), *Experience, variation and generalization: Learning a first language (Trends in Language Acquisition Research Series)*, 167–178. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/tilar.7.10arn>
- Baker, Mark, Kyle Johnson & Ian Roberts. 1989. Passive arguments raised. *Linguistic Inquiry* 20(2). 219–251.
- Bannard, Colin & Danielle Matthews. 2008. Stored word sequences in language learning: The effect of familiarity on children’s repetition of four-word combinations. *Psychological Science* 19(3). 241–248. DOI: <https://doi.org/10.1111/j.1467-9280.2008.02075.x>
- Bargmann, Sascha & Manfred Sailer. 2015. The syntactic flexibility of non-decomposable idioms. In *Posters presented at the PARSEME 4th general meeting*, 19–20.
- Bonin, Patrick, Alain Méot & Aurelia Bugaiska. 2013. Norms and comprehension times for 305 French idiomatic expressions. *Behavior Research Methods* 45(4). 1259–1271. DOI: <https://doi.org/10.3758/s13428-013-0331-4>
- Brysbaert, Marc & Michael J. Cortese. 2010. Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *The Quarterly Journal of Experimental Psychology* 64(3). 545–559. DOI: <https://doi.org/10.1080/17470218.2010.503374>
- Cacciari, Cristina & Maria C. Levorato. 1989. How children understand idioms in discourse. *Journal of Child Language* 16(2). 387–405. DOI: <https://doi.org/10.1017/S0305000900010473>
- Caillies, Stéphanie & Sandrine Le Sourn-Bissaoui. 2006. Idiom comprehension in French children: A cock-and-bull story. *European Journal of Developmental Psychology* 3(2). 189–206. DOI: <https://doi.org/10.1080/17405620500412325>
- Cain, Kate, Andrea S. Towse & Rachael S. Knight. 2008. The development of idiom comprehension: An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology* 102(3). 280–298. DOI: <https://doi.org/10.1016/j.jecp.2008.08.001>
- Clark, Eve V. & Barbara F. Hecht. 1983. Comprehension, production, and language acquisition. *Annual Review of Psychology* 34(1). 325–349. DOI: <https://doi.org/10.1146/annurev.ps.34.020183.001545>
- Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8(2). 81–120. DOI: <https://doi.org/10.1111/j.1467-9612.2005.00076.x>
- Dubinsky, Stanley & Silvester R. Simango. 1996. Passive and stative in Chichewa: Evidence for modular distinctions in grammar. *Language* 72(4). 749–781. DOI: <https://doi.org/10.2307/416101>



- Everaert, Martin. 2010. The lexical encoding of idioms. In Malka Rappaport-Hovav, Edit Doron & Ivy Sichel (eds.), *Lexical semantics, syntax, and event structure*, 76–98. Oxford: OUP. DOI: <https://doi.org/10.1093/acprof:oso/9780199544325.003.0005>
- Fadlon, Julie, Julia Horvath & Tal Siloni. 2014. Decomposability and idiom retrieval: An experimental study. Ms. Tel Aviv University.
- Fadlon, Julie, Julia Horvath, Tal Siloni & Kenneth Wexler. To appear. *Idiom storage and the verbal passive: No unique phrasal idioms*. Language Science Press.
- Gibbs, Raymond W. 1987. Linguistic factors in children's understanding of idioms. *Journal of Child Language* 14(3). 569–586. DOI: <https://doi.org/10.1017/S0305000900010291>
- Gibbs, Raymond W. 1991. Semantic analyzability in children's understanding of idioms. *Journal of Speech, Language, and Hearing Research* 34(3). 613–620. DOI: <https://doi.org/10.1044/jshr.3403.613>
- Gibbs, Raymond W., Nandini P. Nayak & Cooper Cutting. 1989. How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language* 28(5). 576–593. DOI: [https://doi.org/10.1016/0749-596X\(89\)90014-4](https://doi.org/10.1016/0749-596X(89)90014-4)
- Horvath, Julia & Tal Siloni. 2008. Active lexicon: Adjectival and verbal passives. In Sharon Armon-Lotem, Gabi Danon & Susan D. Rothstein (eds.), *Current issues in generative hebrew linguistics*, 105–134. Amsterdam: John Benjamins.
- Horvath, Julia & Tal Siloni. 2009. Hebrew idioms: The organization of the lexical component. *Brill's Journal of Afroasiatic Languages and Linguistics* 1(1). 283–310. DOI: <https://doi.org/10.1163/187666309X12491131130666>
- Horvath, Julia & Tal Siloni. 2016. *Idioms: The type-sensitive storage model*. Submitted.
- Laval, Virginie. 2003. Idiom comprehension and metapragmatic knowledge in French children. *Journal of Pragmatics* 35(5). 723–739. DOI: [https://doi.org/10.1016/S0378-2166\(02\)00137-6](https://doi.org/10.1016/S0378-2166(02)00137-6)
- Leverato, Maria C., Barbara Nesi & Cristina Cacciari. 2004. Reading comprehension and understanding idiomatic expressions: A developmental study. *Brain and Language* 91(3). 303–314. DOI: <https://doi.org/10.1016/j.bandl.2004.04.002>
- Leverato, Maria C. & Cristina Cacciari. 1992. Children's comprehension and production of idioms: The role of context and familiarity. *Journal of Child Language* 19(2). 415–433. DOI: <https://doi.org/10.1017/S0305000900011478>
- Leverato, Maria C. & Cristina Cacciari. 1995. The effects of different tasks on the comprehension and production of idioms in children. *Journal of Experimental Child Psychology* 60(2). 261–283. DOI: <https://doi.org/10.1006/jecp.1995.1041>
- Leverato, Maria C. & Cristina Cacciari. 1999. Idiom comprehension in children: Are the effects of semantic analysability and context separable? *European Journal of Cognitive Psychology* 11(1). 51–66. DOI: <https://doi.org/10.1080/713752299>
- Libben, Maya R. & Debra A. Titone. 2008. The multidetermined nature of idiom processing. *Memory & Cognition* 36. 1103–1121. DOI: <https://doi.org/10.3758/MC.36.6.1103>
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2). 14–25.
- Meltzer-Asscher, Aya. 2012. Verbal passives in English and Hebrew: A comparative study. In Martin Everaert, Marijana Marelj & Tal Siloni (eds.), *The Theta System: Argument structure at the interface*, 279–307. Oxford: OUP. DOI: <https://doi.org/10.1093/acprof:oso/9780199602513.003.0011>
- Mishani-Uval, Yael & Tal Siloni. 2016. Ditransitive Idioms in Hebrew. *Natural Language and Linguistic Theory* 35(3). 715–749. DOI: <https://doi.org/10.1007/s11049-016-9354-8>



- Nippold, Marilyn A. & Catherine L. Taylor. 1995. Idiom understanding in youth: Further examination of familiarity and transparency. *Journal of Speech, Language, and Hearing Research* 38(2). 426–433. DOI: <https://doi.org/10.1044/jshr.3802.426>
- Nippold, Marilyn A. & Catherine L. Taylor. 2002. Judgments of idiom familiarity and transparency: A comparison of children and adolescents. *Journal of Speech, Language, and Hearing Research* 45(2). 384–391. DOI: [https://doi.org/10.1044/1092-4388\(2002/030\)](https://doi.org/10.1044/1092-4388(2002/030))
- Nippold, Marilyn A. & Mishelle Rudzinski. 1993. Familiarity and transparency in idiom explanation: A developmental study of children and adolescents. *Journal of Speech, Language, and Hearing Research* 36(4). 728–737. DOI: <https://doi.org/10.1044/jshr.3604.728>
- Nippold, Marilyn A. & Stephanie T. Martin. 1989. Idiom interpretation in isolation versus context: A developmental study with adolescents. *Journal of Speech, Language, and Hearing Research* 32(1). 59–66. DOI: <https://doi.org/10.1044/jshr.3201.59>
- Nordmann, Emily & Antonia A. Jambazova. 2017. Normative data for idiomatic expressions. *Behavior Research Methods* 49(1). 198–215. DOI: <https://doi.org/10.3758/s13428-016-0705-5>
- Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language*. 491–538.
- Riehemann, Susanne Z. A. 2001. *A constructional approach to idioms and word formation*. Stanford, CA: Stanford University dissertation.
- Ruwet, Nicolas. 1991. *Syntax and human experience*. Chicago, IL: University of Chicago Press.
- Schenk, Andre. 1995. The syntactic behavior of idioms. In Martin Everaert, Erik-Jan van der Linden, Andre Schenk & Rob Schreuder (eds.), *Idioms: Structural and psychological perspectives*, 253–272. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Titone, Debra A. & Cynthia M. Connine. 1994. Comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20(5). 1126–1138. DOI: <https://doi.org/10.1037/0278-7393.20.5.1126>
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- van de Voort, Marlies E. & Wietske Vonk. 1994. You don't die immediately when you kick an empty bucket: A processing view on semantic and syntactic characteristics of idioms. In Martin Everaert, Erik-Jan van der Linden, Andre Schenk & Rob Schreuder (eds.), *Idioms: Structural and psychological perspectives*, 283–300. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Vosniadou, Stella. 1987. Children and metaphors. *Child Development* 58(3). 870–885. DOI: <https://doi.org/10.2307/1130223>
- Webelhuth, Gert & Farrell Ackerman. 1999. A lexical-functional analysis of predicate topicalization in German. *Journal of Germanic Linguistics* 11(1). 1–61. DOI: <https://doi.org/10.1017/S1040820700002456>
- Winner, Ellen. 1988. *The point of words: Children's understanding of metaphor and irony*. Cambridge, MA: Harvard University Press.
- Wulff, Stefanie. 2010. Marrying cognitive-linguistic theory and corpus-based methods: On the compositionality of English V NP-idioms. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: Corpus-driven approaches*, 223–238. Berlin/New York: Walter de Gruyter.

**How to cite this article:** Fadlon, Julie, Julia Horvath, Tal Siloni and Kenneth Wexler. 2018. The acquisition of Hebrew idioms: Stages, internal composition, and implications for storage. *Glossa: a journal of general linguistics* 3(1): 99.1–25, DOI: <https://doi.org/10.5334/gjgl.542>

**Submitted:** 08 October 2017    **Accepted:** 17 May 2018    **Published:** 11 September 2018

**Copyright:** © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS The Open Access logo, which is a stylized 'a' inside a circle.