Idiom Storage and the Mental Lexicon:
Psycholinguistic Reality
Tal Siloni, Julia Horvath, Hadar Klunover and Kenneth Wexler

Abstract
Using an innovative methodology, the paper reports psycholinguistic work that sheds light on the organization of the mental lexicon, the storage technique of phrasal idioms, and the derivation of various diatheses (voices). We conducted a psycholinguistic experiment to examine the storage technique of phrasal idioms. Native speakers of Hebrew were taught invented idioms (in Hebrew) composed on the basis of French idioms. The idioms were headed by predicates of three different diatheses: a verbal passive, an adjectival passive and an unaccusative verb. After learning the idioms, participants had to evaluate for each idiom how likely it is that it shares its idiomatic meaning with its transitive version. The results show that the distribution of phrasal idioms depends on the diathesis of their head. Subjects perceived the likelihood of the verbal passive to share idiomatic meanings with its transitive counterpart as significantly higher than that of both the adjectival passive and the unaccusative. The findings provide support to the following major claims. Phrasal idioms must be stored in the mental lexicon and not in an extra-grammatical component since speakers' conception of them turns out to be sensitive to grammatical features such as diathesis. Moreover, given that the distribution of phrasal idioms is perceived as contingent upon the diathesis of the head, it follows that they must be stored as subentries of the head. In addition, the findings provide support for the hypothesis that adjectival passives and unaccusative verbs must be listed in the mental lexicon, but not the verbal passive. Since the former are derived entries, it follows that the lexicon must be an active component of Grammar, in sharp contrast to claims currently advanced by major trends in theoretical linguistics.

1. Introduction

The human language faculty is assumed by most trends in theoretical linguists to be modular, comprising different, independently operating systems that interface with one another. Theoretical models of the mental grammar assume the existence of a mental lexicon, a syntactic, computational component (the structure building component), a semantic component, and a component responsible for the physical aspect of the language – sound or sign. The modularity of the language faculty has received much support in the past half century of research. Beyond providing additional evidence for the modular architecture of Grammar, this paper concentrates on the lexical component, its nature, and interface with the syntax.

Linguistic research has been concerned with the interface between the mental lexicon and the syntax ever since the inclusion of a lexical component in the theory of grammar (Chomsky 1965). In the past decade, two major interrelated questions have been at the heart of far-reaching controversies on the topic. First, the Word vs. Root question: what kind of entries make up the mental lexicon? Is it words, as was
recently argued again by Horvath and Siloni (2009a)? Or is it more basic elements such as roots, whether consonantal roots as in Semitic languages (Arad, 2005; Doron 2003) or roots in the wider sense, as argued for other languages (e.g., by Borer 2005; Marantz 1997; Pylkkänen 2002 and Ramchand 2006)? The second question concerns the workings of the mental lexicon. Words often seem to derive from other words or from more basic units, for example, the Hebrew verb *hupal* ‘fall (verbal passive)’ may derive from its transitive counterpart, *hipil* ‘fall (transitive)’, or possibly from the consonantal root: n.f.l. Is the mental lexicon an active system that involves diathesis (voice) derivation? Or does it comprise only lists of underived (atomic) entries, while the derivation operations all occur in the syntax? Since under modular approaches, the lexicon regularly feeds the syntax with lexical items, but not vice versa, if the lexicon turns out to involve derived diatheses, then, obviously, the operations deriving them must be lexical. But if the lexicon is root-based, then there may be no reason whatsoever to assume that the lexicon is operative (active); in particular, diatheses derivation may be entirely syntactic, as currently argued by syntacticocentric approaches (Borer 2005; Marantz 1997; Pylkkänen 2002; Ramchand 2006, Harley and Noyer 1999, among many others).

Bat-El (1994), Bolozky (1978), Horvath (1981), and Ussishkin (1999) have argued, based on morphological evidence, that the basic unit of the lexicon is the word (and not the root). These studies show that the input to certain morphological processes is necessarily a word and not a root. Assuming that these morphological processes occur in the mental lexicon, they conclude that the mental lexicon must contain words. However, the assumption that morphological processes occur in the lexicon is no longer a consensus: Recent models of grammar assume that words are associated with their morpho-phonological "clothing" only after the syntactic derivation (e.g. Marantz 1993). If so, then the conclusion that morphological processes can take words as their inputs does not prove that words are listed in the lexicon.

This study deals with the thematic derivation of verbs and adjectives, that is, the derivation setting their semantic and syntactic valence, their diathesis (voice), and not their morpho-phonological derivation. It reports a psycholinguistic experiment that tests the effect various diatheses have on speakers' conception of idiom distribution. Idioms, as is well known, exhibit an intriguing duality: On the one hand, they are complex entities whose internal makeup reflects structural properties of phrasal units. On the other hand, their meaning is conventionalized and cannot be predicted based on semantic composition; therefore, it must be stored in the mental representations. Owing to this duality, idioms constitute a unique source of information regarding the structure of the mental lexicon, its storage techniques, and its interface with the syntax. The present study tests the psycholinguistic reality of the constraints diathesis imposes on idiom storage and consequently distribution.

What we mean by ‘psycholinguistic reality’ is the following. The set of idioms in a language is fixed the way the set of vocabulary items is. Of course, the set can be expanded, but it is not productive in the way syntactic derivations are. Indeed, most studies of idioms rely on data concerning the idioms that actually exist in the language. A potentially stronger kind of evidence, however, could be provided, if humans can

---

1 For earlier versions of this view of the lexicon, see Aronoff (1976) and Jackendoff (1975), among others.
judge what would be an acceptable pattern of idiom in a given language and what would not be. A major claim of our study is that the human linguistic capacity predicts particular patterns of idiom distribution: certain patterns exist and other patterns do not. The paper reports experimental work testing subjects’ decisions with regard to the patterns of productivity of idioms. We teach participants invented idioms in their native language and test them on these new idioms, aiming to shed light on the pattern of productivity of each idiom type. This is what we mean by a test of the psycholinguistic reality of idioms. That is, there is no potential way that memorization could give a participant the answer. This is particularly important concerning tests of a finite entity like the set of idioms.

The structure of the paper is as follows. Section 2 defines the empirical array (2.1), explains what the relevance of data regarding the distribution of idioms is to the understanding of the structure of the lexicon (2.2), and delineates the current research question (2.3). Section 3 reports the psycholinguistic experiment we conducted on native Hebrew speakers to test their perception of the distribution of idioms, and section 4 discusses the findings.

2. Background

2.1 Defining the empirical array

Following Horvath and Siloni (2009b), we assume that the defining properties of idioms singling them out from other fixed expressions are: conventionality (unpredictability) and figuration (metaphoricity). First, idiom meaning is conventionalized. That is, the meaning of these fixed combinations of words cannot be predicted based on the meaning of their subparts. Second, their meaning must be metaphoric (figurative), not literal. Speakers are often unaware of the rationale behind the metaphoric interpretation, but they do perceive that some metaphoric transfer is involved (as noted by Nunberg, Sag and Wasow 1994). The combination of conventionality and metaphoricity yields idioms. Thus, while (1a) is an idiom, as its meaning is both unpredictable and figurative, (1b) is a proverb: on a par with idioms, it is a fixed expression, but its meaning is neither unpredictable nor metaphoric.

\[(1) \begin{align*}
    a. & \quad \text{The grass is always greener on the other side of the fence} \\
    b. & \quad \text{Absence makes the heart grow fonder.}
\end{align*}\]

The linguistic literature distinguishes between clausal idioms (such as (1a) and (2a)) and phrasal idioms (2b). Clausal idioms include clausal structure as part of their fixed material (e.g., fixed tense, modals, negation, and complementizers), for instance, \textit{can’t} in (2a). In phrasal idioms, in contrast, the fixed combination of words, e.g., \textit{spill the}

\[\text{\underline{2 Additional characteristics typical of idioms have been discussed in the linguistic literature. Nunberg, Sag and Wasow (1994) mention the following tendencies idioms show: (i) inflexibility, that is, a tendency to appear in a limited number of syntactic constructions, (ii) proverbiality: a tendency to describe situations of common social interest, (iii) affect: a tendency to imply evaluation or affective stance towards what the idiom describes, (iv) informality: a tendency to belong to informal or colloquial registers. These are indeed often attested in idioms, but they are not necessary and in addition, they are also widely associated with other conventional expressions (e.g., proverbs and aphorisms).}}\]
beans in (2b), does not involve sentential functional material (Horvath and Siloni 2012).

(2) a. can't see the forest for the trees
b. spill the beans

Horvath and Siloni show that clausal idioms differ substantially from phrasal idioms. In section (2.3), once idiom storage is discussed, we resume discussion of clausal idioms and explain why the present study focuses on phrasal idioms.

The most basic question about idiom storage is whether they are stored as part of our linguistic knowledge in the mental lexicon or as part of our general, non-linguistic information, like historical or geographical facts. The choice between these two basic approaches is quite uncontroversial. Knowledge of Idioms involves form-meaning associations, which is linguistic in nature, and distinct from knowledge of facts of history and other language-independent knowledge. As such, it should a priori be stored in the mental lexicon, where linguistic information is stored. Section 2.3 and the findings of the present research provide further support for the claim that idioms are stored with linguistic knowledge. The next question then is how they are stored within the mental lexicon.

2.2 Diathesis distribution and the head-based storage hypothesis

Horvath and Siloni (2009a) present a quantitative corpus-based study that supports the storage hypothesis (3), which they put forth for phrasal idioms:

(3) The Head Based Storage Hypothesis
Phrasal idioms are stored as subentries of their (lexical) head.

Horvath and Siloni examine two options as to how this lexical head is represented in the lexicon: (a) by its root, or alternatively, (b) by the actual predicate (verb, adjective) heading the idiom. Let us discuss a concrete example. Consider the Hebrew idiom in (4), which appears in both the transitive (4a) and unaccusative diathesis (4b). The storage options are as follows: (i) each version of the idiom is stored under its corresponding verbal realization (diathesis), or alternatively (ii) the lexical entry is an abstract verbal representation, unspecified for diathesis, such as the consonantal root y.c.a., under which the idiom is stored, and can therefore be realized in any of its diatheses.

(4) a. hoci et X me-ha-kelim
put.out.TRANS OM X from-the-instruments
‘make X lose his temper’
b. yaca me-ha-kelim
got.out.UNACC from-the-instruments
‘lost his temper’

If the idiom is stored as a subentry of the root, a priori, we would expect distribution across diatheses, that is, idiomatic meanings should be shared by the different

3 OM stands for object marker; roughly, the latter appears in Hebrew with definite objects.
diatheses associated with the root, as it represents them all. If, in contrast, the lexical entry under which the idiom is stored is the actual predicate (diathesis) heading it, distribution across diatheses is not dictated: we can expect some idioms to occur with certain diatheses and not with others, as they do not share a single lexical listing. The two alternatives of the head-based storage hypothesis and the predictions they give rise to regarding idiom distribution are summarized below:

(5) The Head-Based Storage Hypothesis: Root versus Word
   a. The root alternative: phrasal idioms are stored as subentries of the root that represents their head in the mental lexicon.
      Prediction: systematic distribution across diatheses is expected (modulo gaps due to independent reasons).
   b. The word alternative: Phrasal idioms are stored as subentries of the predicate (the word) that heads them.
      Prediction: varied idiom distribution across diatheses is expected to be possible.

In a pioneering corpus-based study designed to test the above predictions, Horvath and Siloni (2009a) examined the distribution of four different diatheses in Hebrew phrasal idioms: the transitive diathesis, the verbal passive, the unaccusative, and the adjectival passive. Seven idiom dictionaries were scanned (Avneyon 2002, Cohen 1999, Dayan 2004, Fruchtmann, Ben-Nathan and Shani 2001, Levanon 1981, Rosenthal 2005, Sévenier-Gabriel 2004), Google searches conducted and complemented by native speakers' judgments in order to examine whether phrasal idioms show cross-diatheses distribution or can give rise to "unique idioms". Unique idioms are defined as idioms that a given diathesis does not share with its root-counterpart that it would most reasonably be related to derivationally. (6) gives Horvath and Siloni's explicit definition of uniqueness for the various diatheses examined in their study.

(6) Idiom Uniqueness
   a. An idiom headed by an unaccusative, a verbal passive or an adjectival passive is unique, if the relevant diathesis has a corresponding transitive alternate, but the latter does not share the relevant idiomatic meaning.
   b. An idiom headed by a transitive verb is unique, if the verb has a corresponding unaccusative alternate, but the latter does not share the relevant idiomatic meaning.

60 predicates of each of the four diatheses were sampled from a verb dictionary (Stern 1994), and for each diathesis type, the number of predicates heading one or more unique idioms was counted. The results are presented in Table 1.

4 Unaccusatives were diagnosed by tests detecting underlying objects in Hebrew (for more details, see note 4 and Meltzer and Siloni to appear). The categorial status of the passive forms was determined by tests detecting adjectives in Hebrew (see Horvath and Siloni 2008).
### Table 1 The distribution of unique phrasal idioms across diatheses

<table>
<thead>
<tr>
<th>Verbal Passives</th>
<th>Unaccusatives</th>
<th>Transitives</th>
<th>Adjectival Passives</th>
</tr>
</thead>
<tbody>
<tr>
<td>0/60</td>
<td>21/60</td>
<td>23/60</td>
<td>13/60</td>
</tr>
</tbody>
</table>

Out of the 60 verbal passive predicates, none appeared in unique idioms. The difference between this result and the number of predicates participating in unique idioms in each of the other diatheses (unaccusatives, transitives, adjectival passives) was statistically significant ($\chi^2=23.088$ $p=0.0001$, $\chi^2=26.033$ $p=0.0001$, $\chi^2=12.42$ $p=0.0004$ respectively). The difference between the number of unaccusatives, that of transitives and that of adjectival passives that participated in unique idioms was not significant.

The findings show that idioms are sensitive to grammatical information, specifically, to the diathesis of their head. This sensitivity reinforces the claim that idioms are stored as linguistic knowledge, since otherwise there is no reason for them to be affected by grammatical factors. Moreover, the existence of many unique idioms attests against the root-based lexicon alternative (5a), which predicts a strong tendency for cross-diatheses distribution of phrasal idioms. The findings clearly favor the word-based lexicon approach (5b), which allows idiom storage by diathesis and therefore predicts varied distribution across diatheses.\(^5\)\(^6\)

Let us now see how the specific pattern of idiom distribution across diatheses can be accounted for. If transitive verbs, unaccusative verbs, and adjectival passives are entries listed in the mental lexicon, then it follows that an idiom may be uniquely listed under one of them, resulting in a unique idiom. If the verbal passive, in contrast, is not formed and listed in the lexicon, it cannot head unique idioms at all, because there is no lexical entry whose subentry it could be. Indeed, in the linguistic literature it is common practice to assume that the verbal passive is not formed in the lexical component, but constitutes a post-lexical output (Baker, Johnson and Roberts 1989; Collins 2005; Horvath and Siloni 2008; Meltzer 2012, among others). As for the other diatheses, it is debated where they are formed, but there are studies that argue that transitives are lexical entries, which can be input for diathesis derivation (Horvath and Siloni 2011, Reinhart, 2002), and that unaccusatives and adjectival passives are derived by a lexical operation (see Chierchia 2004; Horvath and Siloni 2011; Levin and Rappaport 1995; Koontz-Garboden 2009; Reinhart 2002, for unaccusatives, and Horvath and Siloni 2008; Levin and Rappaport 1986; Meltzer 2011, for adjectival passives). The findings reported in Horvath and Siloni (2009a) reinforce this claim. Unless unaccusatives, transitives, and adjectival passives were listed as entries in the lexicon, they should not be able to appear in unique idioms, just like the verbal passive.

\(^5\) Following Horvath and Siloni’s terminology, by word-based lexicon, we mean a lexicon that must include words. It can in addition list roots, or code them indirectly, if this turns out to be well motivated on linguistic grounds or given evidence regarding their psychological reality.

\(^6\) It has been suggested that idioms, in addition to being stored as subentry of their head, are also stored as subentries of their other components (e.g., Everaert 2010). This is not directly relevant for the ‘word-root’ debate, and will not be discussed.
The fact that unaccusatives and adjectival passives allow listing of subentries means that these diatheses exist as entries in the lexicon and are not lexically derived anew upon each use. The relation between them and their input has to be captured by a lexical rule, but the latter is not reactivated over and over again, with each use of the diathesis by the speaker. Note that the existence of unique idioms for these diatheses does not mean that all the phrasal idioms they appear in are unique, phrasal idioms can appear in two or more diatheses. This is not at all surprising given that these diatheses are related by a lexical operation. For instance, out of 60 transitive predicates with unaccusative counterparts, 16 appeared in idioms that they shared with their unaccusative alternate (see the pair of idioms in (4), for example).

In sum, Horvath and Siloni's findings support the head-based storage hypothesis (5b), a word-based lexicon, and the claim that transitives, unaccusatives, and adjectival passives are entries in the mental lexicon, unlike the verbal passive.

Clausal idioms differ from phrasal idioms in various respects. As mentioned in section 2.1, they are headed by fixed functional material. Functional elements constitute a closed class, have no descriptive content, and have no thematic relation to their complement. These among other systematic properties distinguishing functional elements from their lexical counterparts have led linguists to assume that they are stored in their own, functional lexicon (e.g., Emonds 2000). If that is indeed so, it seems unreasonable to assume that clausal idioms, which involve descriptive content, just like phrasal idioms, would be stored under their functional heads in the functional lexicon. Nor does it seem plausible to assume that they are stored under their corresponding verbal head, as they may involve none. Based on that and other evidence, Horvath and Siloni (2012) advance the Independent Storage Hypothesis for clausal idioms: Clausal idioms are stored as units in themselves, not as subentries of other lexical entries. This hypothesis makes various predictions, which are shown to be borne out (see cited reference). Important for our purposes here are the following facts. First, clausal idioms constitute a much smaller set than phrasal idioms, which is attributable to the fact that their storage does not rely on existing subentries and devices (see above-cited reference for discussion and empirical support). Second, clausal idioms turn out to be rigid in their diathesis. This is expected if they are listed as autonomous units and not as subentries of their head, which may be lexically related to another head and thus facilitates sharing of idiomatic meaning. Finally, given the way they are stored, it is not surprising that clausal idioms can be unique to the passive diathesis: If they are not subentries of their head, it is irrelevant whether or not the latter is a lexical entry.

The present study concentrates on phrasal idioms and explores the psycholinguistic reality of their storage. It turns out to provide additional evidence in favor of the head-based storage alternative (5b), and for the conception of the lexicon as an active, word-based component.

---

7 For instance, the clausal idioms in (i) does not include a verbal head:

(i) oznayim la-kotel.
ears to-the-wall
‘Don’t reveal secrets’
2.3 Goal and research questions

We set out to examine the storage procedure of phrasal idioms by means of a psycholinguistic experiment. The empirical question we addressed was whether the diathesis heading the idiom affects speakers' evaluation as to whether the corresponding transitive version of this idiom exists (i.e. the corresponding idiom headed by the transitive alternate).

To check that, we designed an idiom teaching experiment overriding the fact that idioms constitute a closed set per language, by using invented idioms. Given that mastering new idioms means storing them the way other idioms are stored, the experiment opened a window on speakers’ linguistic competence in the domain of idioms and their storage. Native speakers of Hebrew were taught invented Hebrew phrasal idioms, which were formed on the basis of existing idioms in French. The idioms were headed by an unaccusative, a verbal passive, or an adjectival passive. After learning and assimilating the new idioms, participants were presented with each idiom again and asked to estimate how likely it seemed to them that the corresponding transitive version of the idiom exists. In this way, we examined for each idiom how much it was conceived to be able to be unique.

Our predictions were as follows. If phrasal idioms are stored as subentries of the root that represents their head in the lexicon, subjects mastering new phrasal idioms should not favor one diathesis over the other regarding the likelihood of the transitive alternate to exist. In other words, since under the root alternative, idioms are predicted to roughly show consistent distribution across diatheses (5a), subjects should in principle conceive of the various diatheses as equally likely to have a transitive counterpart. That is, the specific diathesis should not affect their decision as to whether or not the relevant idiom can be unique.

On the other hand, if phrasal idioms are stored as subentries of the predicate (the specific diathesis) that heads them (5b), then subjects should accept idioms as unique only in case the relevant diathesis is listed as a lexical entry. In case a particular diathesis is a post-lexical output and is not itself listed in the lexicon, the corresponding idiom should not be judged as likely to be unique. If indeed verbal passives are not entries in the lexicon, unlike the other diatheses under examination (see section 2.2), the results should show a significant difference between the former and the latter with regard to their liability to head unique idioms.

3 The Experiment

3.1 Method

3.1.1 Participants Participants included forty three native speakers of Hebrew, graduate students at Tel Aviv University and at the Tel Aviv-Yaffo Academic College, with no knowledge of French. Their age ranged from twenty to thirty five (mean age 27.5).

3.1.2 Materials and design The stimuli consisted of twelve phrasal idioms in Hebrew especially composed for the purpose of the experiment, based on existing idioms in
French. The idioms were sometimes slightly modified to fit the language and the experimental needs, as detailed below. Out of the twelve idioms, four were headed by a verbal passive, four – by an unaccusative, and the remaining four – by an adjectival passive. All idioms were controlled for (i) idioms' defining properties: conventionality and figuration, and (ii) semantic and pragmatic plausibility of the transitive alternate. For each idiom, a concise explanation of its meaning, and a matching example of usage in context were composed. Table (2) includes examples for each idiom type, with a translation of its meaning and example of usage.

Table 2: Examples of stimuli for each condition

<table>
<thead>
<tr>
<th>Diathesis</th>
<th>Idiom</th>
<th>Meaning</th>
<th>Usage Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaccusative</td>
<td>yaca im snapir be-yad-o</td>
<td>ended up with nothing of what was expected</td>
<td>The political candidate invested in a large, grandiose campaign, but eventually, because of sloppy administration, he came out with a fin in his hand: not even passing electoral threshold.</td>
</tr>
<tr>
<td></td>
<td>came.out with fin in-hand-his</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal passive</td>
<td>Nidla mi-mayim axurim</td>
<td>was obtained profiting from a problematic situation</td>
<td>All his political success was extracted from turbid water, starting with his appointment to minister, exploiting the state of war.</td>
</tr>
<tr>
<td></td>
<td>was.extracted from-water turbid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjectival</td>
<td>matu'ax be-arba sikot</td>
<td>dressed up elegantly</td>
<td>Usually his appearance was frumpy and disheveled. However, when he went to the job interview yesterday, he was all stretched with four pins.</td>
</tr>
<tr>
<td>passive</td>
<td>stretched in-four pins</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Four forms were composed. Form I contained the list of idioms in pseudo-random order. Form II had the list of idioms (in the same pseudo-random order), each accompanied by its literal meaning and an example of usage. Form III included two tasks: (i) a completion task: idiom completion based on the beginning of the idiom (including its head). (ii) a multi-choice comprehension question for each idiom in context, as illustrated in (8-9) below for the idiom in (7). Subjects were asked to fill in this form for purposes of practice and to assure the idioms were incorporated in their lexicon.

---

8 Standard diagnostics were used to determine the diathesis. Intransitives were classified as unaccusatives based on the standard diagnostics for Hebrew (i) licensing of strict verb subject order (Shlonsky 1997, Meltzer and Siloni to appear), (ii) licensing of possessive datives (Borer and Grodzinski 1986, Meltzer and Siloni to appear), (iii) existence of a transitive alternate whose external role is a Cause role (indifferent with regard to the mental state of the argument) (Reinhart 2002). To assure correct identification of the adjectival passive (versus the verbal passive), we used only adjectival passives of the CaCuC template, which is reserved exclusively for adjectival passives.

9 The inspiring French idioms are finir en queue de poisson ‘ended up with nothing of what was expected,’ Pêcher en eau trouble ‘profit from a problematic situation’; être tiré à quatre épingles ‘dressed with lots of care’.
(7) matu'ax be-arba sikot
    stretched in-four pins (that is, stretched with four pins)
    'elegantly, meticulously dressed'

(8) A: what should I wear for the date?
    B: wear something nice, which looks good on you, but on the other hand, don't
    come all stretched with four pins.

(9) In the dialogue above, when B says "stretched with four pins", she means by it that:
    1. A should not come with her hair tied up with four pins
    2. A should not come too well dressed
    3. A should not come stressed out.

Form IV was the target questionnaire: for each of the twelve idioms, subjects had to rank the extent to which they perceived the idiom as likely to exist in its transitive version, on an ordinal scale from 1 (least likely) to 5 (most likely). A translation of the target question corresponding to the idiom in (7) is given in (10). The transitive form was presented in the impersonal, third person plural.

(10) You have learned the idiom "stretched with four pins". How likely (from 1-5) does it seem to you that the following idiom exists as well?

    matxu oto be-arba sikot
    stretched him in-four pins (stretched him with four pins)
    'dressed him elegantly, meticulously'

3.1.3 Procedure The experiment proceeded in two steps. In the first session, Form I (bare idiom list) was handed out to participants. Next, idioms were taught based on Form II: each idiom was read aloud along with its meaning and an example of usage. Participants had only Form I in order to direct their attention to the experimenter's explanation. Understanding was verified. At the end of the session, Form II was handed out to participants, who were asked to study the idioms again at their leisure before the second session. The second session took place 2-3 days later. Idioms were reviewed, by reading them aloud again one by one alongside their meanings. Next, Form III (the testing form), which diagnosed idiom assimilation, was handed out and participants were asked to fill in the completion tasks and answer the multi-choice questions. Immediately afterwards, participants received the target questionnaire (Form IV) and were asked to rate the likelihood that the transitive version of the idiom existed. Data from participants who had errors or missing responses in the testing form were not taken into consideration in the results.

3.2 Results

Fourteen subjects had errors in the testing form and therefore their questionnaires were set aside. The statistical analyses were performed on the remaining twenty nine subjects. A Wilcoxon signed-rank\textsuperscript{10} test comparing participants' median ratings per

\textsuperscript{10} Since ordinal scales may be incompatible with the basic assumptions required for the t test, we preformed the Wilcoxon signed-ranks test.
diathesis revealed the following pattern. Subjects ranked idioms headed by the verbal passive (median=4.5) as significantly more likely to share idiomatic meanings with their transitive counterparts than they ranked idioms headed by both the unaccusative (median=2.5) (two-tailed: W(26)=324, p<0.0001) and the adjectival passive (median =3) (two-tailed: W(25)=296, p=0.0001). The likelihood of the unaccusative and the adjectival passive to share their idiomatic meanings with their transitive alternate was not significantly different (two-tailed: W(26)=142, p=0.072).

![Chart 1 Median score for each diathesis](image)

For each diathesis, the observed patterns of median scores per participant were compared to 3, the hypothetical median score expected if responses were given randomly, with no guiding criterion. A Wilcoxon signed-rank test revealed the following patterns. The verbal passive median scores were significantly higher than the hypothetical median score (W(28)= 314 p<0.0004 2-tailed). The unaccusative median scores were not significantly different from the hypothetical median score (W(29)= 158 p=0.09 2-tailed). The adjectival passive median scores were not significantly different from 3 (W(19)=13 p=0.8 2-tailed) either.
Finally, the patterns of median scores per participant observed for the verbal passive were compared to 5, the highest possible score. A Wilcoxon signed-ranked test shows that the verbal passive median scores (4.5) were significantly lower than 5 (two-tailed W(17)= 153, p=0.0003).

4. Discussion

The experiment reveals that there is a significant difference between speakers' conception of idioms headed by a verbal passive versus their conception of idioms headed by either an unaccusative or an adjectival passive. Speakers perceive idioms headed by the verbal passive as significantly more likely to share idiomatic meanings with their transitive alternates compared to idioms headed by either the unaccusative or the adjectival passive. This kind of difference was not found between idioms headed by an unaccusative versus those headed by an adjectival passive.

The results provide support for the claim that idioms are stored in the mental lexicon, where linguistic knowledge is stored. If idioms were stored as part of general knowledge with other memorized data, speakers' perception of their distribution should not be sensitive to grammatical features such as the diathesis of the predicate. Since speakers do show such sensitivity, it follows that idioms are stored in the mental lexicon together with other form-meaning associations.

The results also reinforce the claim that phrasal idioms are stored as subentries of their lexical head, as suggested by the head-based storage hypothesis, and not as independent entries of their own (unlike clausal idioms). If phrasal idioms were not stored as subentries, there would be no reason for speakers to treat them differently depending on the diathesis of their head.

More specifically, the results corroborate the head-based storage alternative (5b), according to which phrasal idioms are stored as subentries of the actual predicate (diathesis) heading them. The root-storage option turns out to be implausible, because if phrasal idioms were stored under the root of their head, all other things being equal,
we would expect the three diatheses to be conceived by speakers as equally likely to share idiomatic meanings with their transitive alternates. Our results provide psycholinguistic support to the findings reported in Horvath and Siloni's (2009a) corpus-based study.

The strong tendency of verbal passives to share idiomatic meanings with their transitive counterparts (or in other words, the low tendency of the verbal passive to participate in unique idioms) supports the assumption that verbal passives are not entries in the lexicon and are formed post-lexically, as has been claimed by Baker, Johnson and Roberts (1989), Collins (2005), Horvath and Siloni (2008), Meltzer (2012) on independent grounds. This also demonstrates the inaccessibility of predicates formed post-lexically to the lexical component, as predicted by the modular approach to Grammar.

Idioms headed by an unaccusative and those headed by an adjectival passive were conceived as significantly more likely than the verbal passive to appear in unique idioms. This reinforces the claim that both these diatheses are listed in the lexicon. No significant difference was found between speaker's conception of unaccusative and adjectival passive idioms regarding uniqueness. Indeed, if both are lexical entries, there should not be a difference between them regarding the tendency to occur in unique idioms.

Adjectival passives are standardly taken to be derived entries; unaccusatives are taken to be derived by many. Given a modular conception of Grammar, where lexical entries are inserted into the syntactic component, but not vice versa, it follows that the derivation of adjectival passives and unaccusatives must be lexical, as they exist in the lexicon. This has been independently suggested by Chierchia (2004), Horvath and Siloni (2011), Koontz-Garboden (2009) Levin and Rappaport (1995), Reinhart (2002) for unaccusatives, and by Horvath and Siloni (2008), Levin and Rappaport (1986), Meltzer (2011), Wasow (1977) for adjectival passives. If these diatheses are derived by lexical operations, it means that the lexicon must be an active component of Grammar (as argued by Reinhart 2002, Siloni 2002 among others), and not a mere storehouse of atomic items, in contrast with the view of syntacticocentric approaches (Borer 2005; Marantz 1997; Pylkkänen 2002; Ramchand 2006, Harley and Noyer 1999, among many others).

Moreover, as there is evidence that unaccusatives and adjectival passives are entries in the lexicon, it follows that they are not lexically derived anew upon each retrieval. That is, the organization of the lexical component should, of course, capture the relation between derived diatheses and their input, but the derivation itself does not apply with each use of the diathesis by the speaker, as the diathesis itself is listed.11

The claim that unaccusatives and adjectival passive are lexical entries receives further support from the comparison of their observed patterns of median scores per participant with 3, the hypothetical median score expected if responses were given randomly, with no systematic guiding criterion (see chart 2). Let us see why.

11 The operations responsible for lexical derivations are undoubtedly available (despite the fact that they are not activated for each use), not only at the acquisition stage for formation of new lexical items, but also beyond this stage, as attested by the fact that they can apply to neologisms, loanwords and invented items.
If indeed unaccusatives and adjectival passives are entries in the lexicon, idioms they head may, but do not have to, be available to their transitive alternate. Horvath and Siloni’s corpus-study has found, out of 60 transitive-unaccusative pairs, 21 unique idioms headed by an unaccusative, 23 unique idioms headed by a transitive, and 16 idioms that unaccusatives and their transitive alternates share. There does not seem to be any criterion systematically determining when an idiomatic meaning is available to both alternates and when it is unique to one diathesis. For example, there is no obvious reason why (11a) is unique while (4b) repeated as (11b) exists also in the transitive diathesis (see (4a)).

(11)  a. yaca be-šen va-ayin
dent.out in-tooth and-eye’
‘ended up losing much’
b. yaca me-ha-kelim
dent.out.UNACC from-the-instruments
‘lost his temper’

This means that subjects had no criterion for deciding whether or not an unaccusative idiom and an adjectival passive one were unique. Their responses thus could be either randomly distributed across the scale values (chance distribution of hypothetically 20% per value), or more condensed towards the median owing to avoidance of edge values triggered by lack of basis for decision (a “to be on the safe side” option). Importantly for our purposes, under both options the responses should be equally distribution around 3, thus, yielding the hypothetical median 3.

Wilcoxon signed-rank tests revealed that both the unaccusative median scores and the adjectival passive median scores were not significantly different from the hypothetical median score, 3. This is expected since participants had no criterion for decision.

The Wilcoxon test also shows that the verbal passive median scores were significantly higher than the hypothetical median score. This shows that indeed in this case a systematic principle directed the responses in one direction. More specifically, it shows that subjects clearly tend to perceive verbal passive idioms as non-unique, in contrast with their perception of adjectival passives and unaccusatives.

The following question then arises: If the verbal passive cannot give rise to unique phrasal idioms since it is not listed in the lexicon, why did we not get absolute results but only strong statistical tendencies? Recall that the median scores per participant observed for the verbal passive were found to be significantly lower than the highest possible score. That is, why is it not the case that participants always give the maximum likelihood score (5) to the question regarding the existence of a transitive counterpart to the verbal passive idiom?

We did not expect absolute results. First, the questionnaire offered a scale of 1 to 5 and not yes/no decisions. Since the subjects had a scale at their disposal, it seems not surprising that they made use of it and did not opt for absolute responses, the participants' rationale being that scaled options would not be put at their disposal unless they were expected to make use of the scale. More importantly, clausal idioms often look superficially just like verb phrase idioms, as illustrated in (12). However,
 unlike phrasal idioms (e.g. (12a)), their tense turns out to be fixed upon further examination: (12b) does not occur in the present and future tenses, and is therefore classified as clausal, as its tense is fixed.

(12) a. yaca/yoce/yece me-ha-kelim
go.out.PAST/PRES./FUT. from-the-instruments
‘lost/loses/will.lose/ his temper’

b. yaca/*yoce/*yece ve-ozn-av mekutafot
went.out and-ears-his torn.out
‘failed and was ashamed and disappointed’

As mentioned in section 2.2, clausal idioms are unique to their diathesis, including the verbal passive diathesis, as illustrated in (13).

(13) a. hutla (*mutlet/*total) ha-kubiya
cast.PASSIVE.PAST(PRES./FUT.) the-dice
‘The die was cast.’

b. hetilu et ha-kubiya.
throw. ACTIVE.PAST ACC the-dice
‘Threw. PASSIVE. PAST.IMPER. the dice’ (no idiomatic meaning) 12

Since participants received no information whatsoever as to idiom availability in other tenses, and therefore could have unconsciously consider them occasionally as clausal, we have all the more reason to expect non-absolute responses. 13 Nonetheless, the strong statistical tendency the verbal passive showed to be non-unique (in comparison to the patterning of the unaccusative and adjectival passive diatheses) was expected. To see why, recall that clausal idioms are much rarer than verb phrase idioms. Idioms that are clausal only due to fixed tense are even rarer, as they constitute a subset of the set of clausal idioms. Therefore, in light of their low frequency, we do expect speakers to opt for the more probable hypothesis, namely, that the idioms they acquired were phrasal idioms, and accordingly associate the verbal passive ones with transitive alternates.

6. Conclusion

The paper reports a psycholinguistic experiment allowing a glimpse into speakers’ linguistic competence in the domain of idioms and their storage. The experiment was designed to evaluate speakers' conception of the pattern of distribution of phrasal idioms. The results contribute to our understanding of the nature of the mental lexicon, the storage of phrasal idioms, and the derivations of several diatheses. First, they provide support that the mental lexicon lists words and is an active component of Grammar, where valence changing operations can apply. As to the storage of phrasal idioms, given that their distribution has turned out to be sensitive to grammatical features, they must be stored in the mental lexicon, and not in an extra-grammatical component. Since the diathesis of the head crucially affects speakers' conception,

12 Google searches yield no idiomatic occurrence.
13 We suspect information regarding fixed tense in idiomatic expressions is hard to assimilate without long term exposure in context.
phrasal idioms must be stored as subentries of their actual head. Finally, the paper offers evidence regarding linguistic controversies as to the locus of derivation of certain diatheses. It strongly suggests that the adjectival passive and the unaccusative are listed in the lexicon, and must therefore be derived by lexical operations, contra syntacticocentric approaches. The verbal passive is distinct: the findings reinforce the view that it is a post-lexical output, not listed in the lexical component, in concert with common approaches.

References:

Horvath, Julia & Tal Siloni. 2009b. *Idioms: Mental Representation and Acquisition*. Unpublished ms., Tel Aviv University.


