

## Idioms: The Type-Sensitive Storage Model

Julia Horvath<sup>1</sup> and Tal Siloni<sup>1,2</sup>

<sup>1</sup>Tel Aviv University and <sup>2</sup>UCL

### 1. Introduction

#### 1.1 Definition of empirical array

- (1) Idioms (core set)  
 Fixed multilexemic expressions whose meaning is  
 a. unpredictable (conventionalized) and  
 b. figurative (metaphoric)

Multilexemic expressions involve special meanings of lexical items relative to one another, and raise questions that the storage of special meanings of single lexical items does not.

Unpredictability Although the level of opacity (or transparency) of idiom meaning changes from one idiom to another (2a-b), the meaning cannot be calculated in advance.

- (2) a. *land on one's feet*  
       ‘make a quick recovery’  
 b. *cool one's heels*  
       ‘wait’

Figuration (3) is not fully predictable but it is not figurative. We focus on fixed expressions involving a figurative interpretation.<sup>1</sup>

- (3) center-divider  
       ‘median strip’

#### 1.2 The Type-Sensitive Storage Model: Phrasal vs. Clausal Idioms

Phrasal Idioms: headed by a lexical head. The idioms in (2) are VP idioms.

Clausal Idioms: headed by a clausal functional head (4); not necessarily full clauses

- (4) a. can't see the forest for the trees                   (modal, negation)  
       ‘doesn't perceive the whole situation clearly due to focusing on the details’  
 b. butter wouldn't melt in X's mouth                   (modal, negation)  
       ‘someone is acting as innocent’  
 c. not have a leg to stand on                           (negation)  
       ‘have no support (for your position)’  
 d. the squeaky wheel gets the grease                   (tense)  
       ‘the most noticeable (loudest) ones are the most likely to get attention’  
 e. where's the beef?                                   (wh-phrase, (contracted) copula)  
       ‘where is the important content?’

---

<sup>1</sup> By figurative interpretation we roughly mean that the event or state (most often concrete) described by the expression is understood as representative of some other (mostly, more abstract) event or state.

Criteria for clausality of idioms: Fixed tense, mood, modal, obligatoriness or impossibility of negation, CP-material, such as a complementizer or a *wh*-phrase.

Idioms exhibit an inherent duality

- Complex entities whose internal makeup reflects structural properties of phrasal units
- Conventionalized meaning that cannot be predicted by semantic composition

→ Idioms must be stored. Since they have syntactic structure, the question as to how they are stored is of particular interest and may shed light on the lexical component.

(5) The Type-Sensitive Storage Model (TSS)

- a. Idioms are stored as part of our linguistic knowledge (not as general, non-linguistic information).
- b. Phrasal idioms – Subentry Storage: Phrasal idioms are stored as subentries of some other existing lexical entry/ies representing their subconstituent(s) in the lexicon.
- c. Clausal Idioms – Independent Storage: Clausal idioms are stored as independent entries on their own, as single units (“big units”).

## 2. Phrasal Idioms

### 2.1 Subentry Storage vs. Independent Storage

Evidence for Subentry Storage

We will present evidence that phrasal idioms are stored as subentries of another, existing entry, not as independent entries of their own.

Subentry Storage

(6) Head Based Storage (head only) vs. Multiple Storage (section 2.3)

- a. Head Based Storage: as subentries of the lexical entry of their head only (Horvath & Siloni 2009)
- b. Multiple Storage: as subentries of the entries of their lexical constituents (head and others) (Everaert 2010, Harley & Noyer 1999)

(7) Root vs. Predicate

- a. Root: as subentries of the root of the head
- b. Predicate: as subentries of the predicate (Horvath & Siloni 2009)

Notice: Subentry storage is contingent upon the listing – existence in the lexicon – of the (mother) entry

## 2.2 Verb phrase idiom surveys in Hebrew and English

Evidence for Subentry storage under predicates (Horvath & Siloni's 2009, 2014 surveys)

Aim: Examine the sporadic observations in the literature wrt the unavailability of idioms specific to the verbal passive (Dubinsky and Simango 1996, Marantz 1997, Ruwet 1991), and compare the distribution of these idioms to that of idioms in other diatheses/voices.

Data and procedure: A random sample of 60 predicates of the following 4 diatheses was collected in Hebrew and in English: unaccusatives, transitives, verbal passives and adjectival passives. The distribution of phrasal idioms across these diatheses was examined in idiom dictionaries (followed by Google searches complemented by native speaker's judgments). The number of predicates of each type appearing in idioms *unique* to the diathesis was counted.

- (8) Uniqueness (definition): An idiom is unique to a given diathesis  $\alpha$ , if  $\alpha$  does not share the idiom with its (existing) root-counterpart  $\beta$ , which  $\alpha$  would most directly be related to by derivation. Specifically,
- a. An idiom in the verbal passive / adjectival passive / unaccusative is unique if the corresponding (existing) transitive does not share the idiom.
  - b. An idiom in the transitive is unique if the corresponding (existing) unaccusative does not share the idiom.

Examples (# means idiom does not exist)

### Unique unaccusative

- (9) a. burst at the seams  
'be extremely full or crowded'  
b. #burst something at the seams

### Unique adjectival passive

- (10) a. bent out of shape  
'be angry, insulted'  
b. #bend someone out of shape

### Unique transitive

- (11) a. turn the tables on X  
'reverse someone's plan'  
b. #The tables turn on X

## Results

Table 1 English: The distribution of unique phrasal idioms across diatheses<sup>2</sup>

Verbal passives	Unaccusatives	Transitives	Adjectival passives
0/60	12/60	22/60	13/60

<sup>2</sup> Statistical analysis: The number of phrasal idioms headed by a verbal passive (0) is significantly different from those headed by an unaccusative ( $\chi^2 = 11.204$ ,  $p = .0008$ ), a transitive ( $\chi^2 = 24.546$ ,  $p < .0001$ ) and an adjectival passive ( $\chi^2 = 12.42$ ,  $p = .0004$ ). The difference between the number of unaccusatives, that of transitives and that of adjectival passives that participated in unique idioms was not significant  $\chi^2(2) = 5.24$ ,  $p = .073$ ).

Table 2 Hebrew: The distribution of unique phrasal idioms across diatheses<sup>3</sup>

Verbal passives	Unaccusatives	Transitives	Adjectival passives
0/60	21/60	23/60	13/60

Verbal passives vs. other diatheses → significant difference

Unaccusatives, transitives, adjectival passives → insignificant difference

In both languages

- Verbal passives cannot have their own, unique phrasal idioms, in contrast to other diatheses. (Phrasal idioms in the verbal passive exist only if the idiomatic meaning is shared by the corresponding transitive (active) alternate.)
- Transitives, unaccusatives and adjectival passives can head unique (as well as shared) phrasal idioms.

## Discussion

**I.** Sensitivity of phrasal idioms to a grammatical factor, the diathesis of their head, attests

- in favor of storage with linguistic knowledge  
Or else why would their distribution depend on the diathesis of the head?
- in favor of subentry storage (against independent storage)  
If phrasal idioms were stored as independent entries on their own, not as subentries of their head, there would be no reason for why their existence is contingent upon the diathesis of the head.

**II.** Pervasive existence of unique phrasal idioms for 3 diatheses provides evidence

- for storage by the predicate and against storage by the root
- that transitives, unaccusatives and adjectival passives are lexical entries
- that verbal passive is different (not a lexical entry/beyond the domain of idiosyncrasy)

Note: Uniqueness and sharing of idioms is not a reflection of productivity/regularity at the diathesis level.

## 2.3 The Multiple Storage Hypothesis?

### (6) Subentry Storage

- a. Head Based Storage: only as subentries of the lexical entry of their head (Horvath & Siloni 2009)
- b. Multiple Storage: as subentries of the entries of their lexical constituents (head and others) (Everaert 2010, Harley & Noyer 1999)

<sup>3</sup> Statistical analysis: The number of phrasal idioms headed by a verbal passive (0) is significantly different from those headed by an unaccusative ( $\chi^2 = 23.088$   $p = 0.0001$ ), a transitive ( $\chi^2 = 26.033$   $p = 0.0001$ ) and an adjectival passive,  $\chi^2 = 12.42$   $p = 0.0004$ . The difference between the number of unaccusatives, that of transitives and that of adjectival passives that participated in unique idioms was not significant  $\chi^2(2) = 4.313$ ,  $p = .116$ .

Horvath & Siloni (2009) advance the Head-based Storage (6a), but the surveys' results (Table 1-2) are also compatible with Multiple Storage (6b)

Initial experimental evidence against multiple storage

Idiom retrieval experiment: a completion task (Efrati, Siloni & Horvath 2011)

Aim: Provide empirical evidence to decide between head-based storage vs. multiple storage

### Retrieval Assumption

Idiom storage is reflected in idiom retrieval. That is, idiom retrieval is facilitated if during retrieval there is access to the lexical entry/ies under which the idiom is stored.

Participants: 80 Native speakers of Hebrew with high school education (minimum)

Method and stimuli: 2 types of questionnaires (see (12-13)) with 18 phrasal idioms of the structure: V NP PP (PP facilitates retrieval) were each filled in by 40 participants to the best of their ability in one minute (time set by pilot trial to impose time pressure).

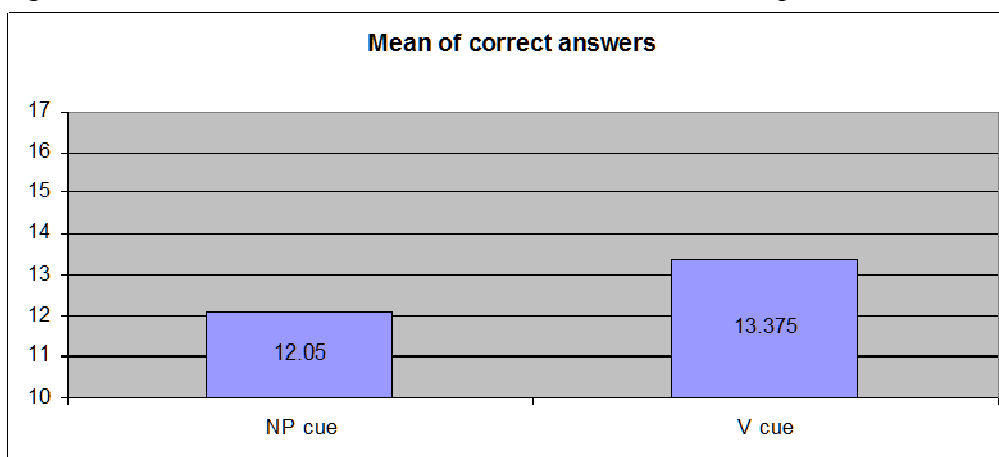
(12) 2 types of questionnaires

- |    |                        |          |            |
|----|------------------------|----------|------------|
| a. | V cue Questionnaire:   | V __ PP  | (blank NP) |
| b. | NP cue Questionnaire : | __ NP PP | (blank V)  |

- (13)a. *sam* \_\_\_\_\_ *al ha-šulxan* (V cue)  
 b. \_\_\_\_\_ *et ha-klafim al ha-šulxan* (NP cue)  
 put OM the-cards on the-table (OM= definite object marker)  
 'revealed one's intentions, put one's cards on the table'

### Results

Figure 1: The mean of correct answers in NP-cue vs. V-cue questionnaires



Statistical analysis: The mean of correct responses for V cue is significantly higher than that of the NP cue (t-test for correlated samples  $t(39)=2.7$ ,  $p=0.01$ ).

## Discussion

V cue questionnaire scored significantly higher. This provides initial evidence

- in favor of our retrieval assumption
- in favor of Head-based storage for phrasal idioms
- against Multiple storage of phrasal idioms
- against Independent Storage of phrasal idioms (as “big” units)

Alternative accounts?

Retrieval of nouns easier than that of verbs? There is no significant difference between the retrieval of nouns and verbs in literal phrases (Ordentlich 2014).

Retrieval of phrase-initial elements harder than that of non-initial ones? There is no significant difference between the retrieval of phrase-initial and non-initial elements in literal phrases (Ordentlich 2014).

- (14) a. The Head-based Storage Hypothesis: Phrasal idioms are stored as subentries of the lexical entry of their head only.
- b. Lexical representation: The head selecting the other idiom constituents by a mechanism also involved in l(exical)-selection (Baltin 1987 suggest that verbs selecting a PP complement utilize l-selection to specify the particular lexical choice of preposition they govern. See also O’Grady 1998, Rothstein 1992).

- (15) a. KICK V <Ag> <Th> (<G>) ‘hit with the foot’  
 b. KICK’ V <Ag > <Th → bucket: N<sub>+DEF</sub> > ‘die’

### 3. Clausal Idioms: Independent (“big unit”) Storage

- (4) a. can't see the forest for the trees (modal, negation)  
 b. butter wouldn't melt in X's mouth (modal, negation)  
 c. not have a leg to stand on (negation)  
 d. the squeaky wheel gets the grease (tense)  
 e. where's the beef? (wh-phrase, (contracted) copula)

Criteria for clausality of idioms: Fixed tense, mood, modal, obligatoriness or impossibility of negation, CP-material, such as a complementizer or a *wh*-phrase.

#### 3.1 The rationale for the Independent Storage hypothesis

- The head of clausal idioms is a *functional*, not *lexical*, element.

(16) Functional elements in contrast to lexical ones

- a. are closed class items
- b. have no descriptive content (Abney 1987)
- c. bear no thematic relation to their complement

Functional elements are argued to be stored in a separate “lexicon” (Emonds’ 2000 “Syntacticon”), or as “f-morphemes”, which are not part of the Encyclopedia (Distributed Morphology).

### Potential storage hypotheses for clausal idioms

- (17) Storage as subentries of their functional head: Storage of clausal idioms, which have descriptive content, as subentries of their functional head (e.g., (4d) as a subentry of present tense morpheme, or (4e) as a subentry of the interrogative complementizer) in the “functional Lexicon”, where entries don’t have descriptive content, **is implausible**. We therefore do not pursue this option.
- (18) Storage as subentries of the lexical head of the “extended projection” (using Grimshaw’s 1991, 2000 term) constituting the clausal idiom, namely, storage as subentries of the verb on a par with VP idioms.
- (19) Independent Storage as an entry (“big unit”) on its own (i.e., not a subentry to another lexical entry)

Independent Storage may be needed for storage of idiomatic expressions that have no recognizable internal structure at all, i.e., have neither a functional nor a lexical head, such as in (20). (Head-based storage option is unavailable for them.)

- (20) a. trip the light fantastic  
‘dance (nimble or lightly)’  
b. happy go lucky  
‘carefree, easy-going’

- (21) Independent Storage: An idiom that is not headed by a lexical category gets stored as a single unit listed as an independent lexical entry.

Subentry storage (18) and independent storage (21) have opposite predictions regarding the distributional properties of clausal idioms: (18), but not (21), predicts the same pattern of sensitivity to diathesis as attested in verb phrase idioms (sections 3.2 and 3.3).

### **3.2 Unique clausal idioms in the verbal passive**

(18) predicts the absence of unique clausal idioms in the verbal passive, on a par with VP idioms.

(21) predicts occurrence of unique clausal idioms in the verbal passive. Under *Independent Storage*, clausal idioms get lexicalized in one piece, following consistent use of the expression in the relevant context. Clausal idioms thus can include outputs of post-lexical operations, and hence, there should be idioms unique to the verbal passive. Initial evidence for *Independent Storage*, to be further tested:

- (22) a. Rome wasn’t built in one day.  
(#They didn’t build Rome in one day.)

- b. may as well be hung/hanged for a sheep as (for) a lamb  
 (#...may as well hang X for a sheep as (for) a lamb)
- c. wouldn't be caught dead...  
 (#...wouldn't catch X dead...)
- d. kurcu me-oto xomer (fixed tense) (Hebrew)  
 made.VPASSIVE of-same material  
 'share the same characteristics'  
 (\*kirec)
- e. nigzezu maxlafot-av (fixed tense)  
 shear.VPASSIVE hair-one's  
 'lost his power/influence'  
 (#gazezu et maxlafot-av: shear OM hair-his)
- f. hutla ha-kubiya (fixed tense)  
 cast. VPASSIVE the-die  
 'The process is past the point of return'  
 (#hetilu et ha-kubiya: cast OM the-die)

### 3.3 Additional evidence for the TSS Model: diathesis rigidity

Phrasal idioms exhibit sharing of idioms across diatheses (Tables 3-4).

#### Examples

##### Shared transitive-unaccusative

- (23) a. burst someone's bubble  
 'destroy someone's illusion'  
 b. someone's bubble burst

##### Shared transitive-adjectival passive

- (24) a. carve something in stone  
 'fix some idea/agreement permanently'  
 b. carved in stone

##### Shared transitive-verbal passive

- (25) a. bring into line  
 'make something like another'  
 b. was brought into line

Table 3 English: The distribution of shared phrasal idioms across diatheses

Transitive- Verbal passive	Transitive- Unaccusative	Transitive- Adjectival passive
35/60	17/60	22/60



Table 4 Hebrew: The distribution of shared phrasal idioms across diatheses

Transitive- Verbal passive	Transitive- Unaccusative	Transitive- Adjectival passive
10/60	16/60	5/60

According to our preliminary investigation clausal idioms, unlike phrasal ones, fail to exhibit sharing across diatheses. They are uniformly unique to their diathesis, irrespective of what this diathesis is, for example:

Clausal idioms: transitive vs. verbal passive: (fixed modal/negation/aux)

- (26) a. can't make heads or tails of X  
'unable to understand X'  
b. #Head or tails can't be made of X. (based on NSW'ex. 1994)
- (27) a. can't see the forest for the trees  
b. #The forest can't be seen for the trees.
- (28) a. could've knocked me over with a feather  
'I was extremely surprised'  
b. #I could've been knocked over with a feather.

Clausal idioms: unaccusative vs. transitive

- (29) a. A journey of a thousand miles begins with a single step.  
'Even the longest and most difficult ventures have a starting point.'  
b. #One begins a journey of a thousand miles with a single step.

Clausal idioms: transitive vs. unaccusative

- (30) a. You can't make an omelet without breaking a few eggs.  
'It is difficult to achieve something important without causing any unpleasant effects.'  
b. #You can't make an omelet without a few eggs breaking.

Clausal idioms: adjectival passive vs. transitive

- (31) a. The road to hell is paved with good intentions.  
People often mean well but do bad things.'  
b. #Good intentions pave the road to hell.

Even when no reordering or addition of words is involved, diathesis alternation does not seem possible (example from Hebrew):

- (32) a. kše-xotvim ecim, nitazim švavim.  
when-chop.TRANS.IMPERS trees, sprinkled.UNACC chips  
'When you act, there are risks.' 'Where trees are felled chips will fly.'  
b. #kše-nextavim ecim, nitazim švavim.  
when-chop.VPASS trees, sprinkled.UNACC chips  
c. #kše-xotvim ecim, metizim švavim.  
when-chop.TRANS.IMPERS trees, sprinkled.TRANS chips

Table 5 summarizes sharing vs. uniqueness of diatheses in phrasal vs. clausal idioms.

	Phrasal	Clausal
<i>Verbal passive</i>	Shared	Unique
<i>Transitive, Unacc., Adjectival passive</i>	shared or unique	Unique

- Sharing of phrasal idioms between the transitive, unaccusative, or adjectival passive is the result of the links between root-related entries in the lexicon. (Sharing is not automatic as it requires additional listing under the relevant entry.)
- Since clausal idioms are stored as independent big entries, not as subentries, sharing of idiomatic meaning across diatheses is not expected.

In sum: Preliminary evidence seems to confirm the Independent Storage proposal for clausal idioms. Their cross-diathesis distribution contrasts with the distribution attested for phrasal idioms, and thus also reinforces our subentry (head-based) storage proposal for phrasal idioms.

### 3.4 Independent Storage: a ‘last resort’ strategy?

The Independent Storage strategy turns out **not** to be a free procedure, but a last resort option. The lack of unique *phrasal* idioms in the verbal passive shows that failure of storage under the verbal form heading the idiom cannot lead to storage as an independent entry. This suggests that independent storage may be more "costly" than the subentry one.

#### Preliminary results

A survey of a sample of 4 letters (2nd to 5<sup>th</sup>) from Rosenthal Hebrew dictionary reveals:  
175 VP idioms vs. 38 clausal ones. (Birger 2014)

## 4. Decomposability

- Decomposable idioms are “idiomatically combining expressions” in Nunberg, Sag and Wasow’s (1994) terms.
- Nondecomposable idioms in their terms are “idiomatic phrases”.  
Following their definition:

### (33) Decomposable vs. nondecomposable idioms: Definition

An idiom is *decomposable* if it is isomorphic with its meaning, in the sense that each of its components corresponds to an element of its meaning; otherwise it is *nondecomposable*.

### 4.1 Phrasal idioms and decomposability

#### Evidence for the split

- Children (age: 6-9) completed nondecomposable phrasal idioms significantly better than decomposable ones.
- Adults complete decomposable phrasal idioms significantly better than nondecomposable ones. (Fadlon, Horvath, Siloni & Wexler (yesterday’s talk))



- (39) a. make sense (out) of X  
 b. (Some) sense has been made out of the proposal.  
 c. ?The proposal has been made sense (out) of. (compare to (26a))
- (40) a. make no bones about X  
 b. No bones were made about the gravity of the situation.
- (41) a. make mincemeat (out) of X  
 b. Mincemeat will be made (out) of the opposing side. (based on NSW 1994)
- (42) a. hixnis et X la-tmuna  
 let.in.TRANS OM X to.the-picture  
 ‘involved someone in’  
 b. huxnas la-tmuna  
 let.in.VPASS to.the-picture
- (43) a. hexzir et ha-šed la-bakbuk.  
 returned.TRANS OM the-demon to.the-bottle  
 ‘managed to gain control on a social phenomenon that erupted again’.  
 b. ha-šed huxzar la-bakbuk.  
 the-demon returned.VPASS to.the-bottle
- (44) a. hikši’ax lib-o  
 hardened.TRANS heart-his  
 ‘got stubborn’  
 b. lib-o hukšax  
 heart-X hardened.VPASS

The distinction between decomposable and nondecomposable idioms does not seem to correspond to distinct storage strategies (subentry storage vs. independent storage, respectively).