Idioms: The Type Sensitive-Storage Model

Julia Horvath¹ and Tal Siloni¹,²
¹Tel Aviv University and ²UCL

1. Introduction

1.1 Definition of empirical array

(1) Idioms (core set)
   Fixed multilexemic expressions whose meaning is
   a. conventionalized-unpredictable 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 The level of opacity (or transparency) of idiom meaning changes from one idiom to another. But the meaning cannot be calculated in advance.

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

Metaphoricity Unpredictability is a property idioms share with other fixed expressions (3). Idioms involve a metaphoric (figurative) interpretation.

(3) center-divider
    ‘median strip’

1.2 Our Proposal: 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).

(4) a. can't see the forest for the trees
   b. butter wouldn't melt in X's mouth
   c. not have a leg to stand on
   d. the squeaky wheel gets the grease
   e. what's eating X?

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

a. **Phrasal idioms** – Subentry Storage: Phrasal idioms are stored as subentries of some other existing lexical entry/ies representing their subconstituent(s) in the lexicon.

b. **Clausal Idioms** – Independent Storage: Clausal idioms are stored as independent entries of their own, as single units (“big words”).

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 (only head) vs. Multiple Storage (section 2.3)

a. **Head Based Storage**: only 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** (root+diathesis): 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 and Siloni’s 2009, 2014 surveys)

Aim: Examine the sporadic observations found in the literature regarding the unavailability of unique passive idioms, and compare their distribution to that of other diatheses.

Data and procedure: A random sample of 60 predicates of various diatheses was collected in Hebrew and in English. The distribution of phrasal idioms was examined in Hebrew and English idiom dictionaries across the following three verbal diatheses: intransitives (unaccusative) having transitive alternates, transitives having unaccusative alternates, and verbal passives and their transitive alternates. The number of predicates of each type giving rise to idioms unique to the diathesis was counted.
(8) **Uniqueness** (definition): An idiom is *unique* to a given diathesis (voice) A, if A does not share the idiom with its root-counterpart B, which A would most directly be related to by derivation. Specifically,

a. Verbal passives and unaccusatives (inchoatives): Unique if there is no corresponding transitive idiom.

b. Transitives: Unique if there is no corresponding unaccusative idiom.

**Results** The verbal passive does not occur in unique idioms, unlike the unaccusative and the transitive diatheses, each of which manifested unique (as well as shared) idioms (Table 1 for Hebrew and Table 2 for English).

**Table 1**: The number of Hebrew predicates allowing unique idioms (out of the 60 predicates)

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

Statistical analysis: The number of phrasal idioms headed by a verbal passive (0) is significantly different from those headed by an unaccusative ($\chi^2 = 23.088$, $p < .0001$), and the transitive ($\chi^2 = 26.033$, $p < .0001$). The number of phrasal idioms headed by an unaccusative is not significantly different from those headed by a transitive ($\chi^2 = 0.036$, $p = .8495$).

**Table 2**: The number of English predicates allowing unique idioms (out of the 60 predicates)

<table>
<thead>
<tr>
<th>Verbal Passives</th>
<th>Unaccusatives</th>
<th>Transitives</th>
</tr>
</thead>
<tbody>
<tr>
<td>0/60</td>
<td>12/60</td>
<td>22/60</td>
</tr>
</tbody>
</table>

Statistical analysis: The number of phrasal idioms headed by a verbal passive (0) is significantly different from those headed by an unaccusative ($\chi^2 = 11.204$, $p = .0008$), and the transitive ($\chi^2 = 24.546$, $p < .0001$). The number of phrasal idioms headed by an unaccusative is not significantly different from those headed by a transitive ($\chi^2 = 3.324$, $p = .0683$).

Verbal passives vs. other diatheses → significant difference
Unaccusatives vs. transitives → insignificant difference

**Discussion**

- In both languages, the pattern of distribution of phrasal idioms is dependent on the diathesis of the idiom’s head.

  This sensitivity constitutes evidence against Independent (“big unit”) Storage for phrasal idioms and in favor of Subentry Storage.

- Verbal passives cannot list 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, under which this meaning is stored). This follows if:
  1. Phrasal idioms are subentries of their predicates.

- If only roots were listed, we would a priori expect shared idioms across the board for all alternates from the same root (unless independently blocked) (for arguments against approaches assuming a syntactic domain for special meaning, see Horvath & Siloni 2013).

**Phrasal idioms must be stored as subentries of predicates, not roots.**

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

If this account is correct, then

(9) The division of labor between the lexicon and syntax:
   a. Verbal passives are formed post-lexically.
   b. Transitives and unaccusatives both are lexical entries (Horvath and Siloni 2011, Meltzer-Asscher 2011, Reinhart 2002).
   c. The lexicon includes derived entries (predicates, not (only) roots), and hence must be an active component of grammar.

### 2.3 The Multiple Storage Hypothesis?

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

Horvath & Siloni (2009) advance the Head-based Storage (6a), but the surveys’ results 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

**Participants** 80 Native speakers of Hebrew with high school education (minimum)
Method and stimuli 2 types of questionnaires (see (10-11)) 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).

(10) 2 types of questionnaires
   a. V cue Questionnaire: V __ PP (blank NP)
   b. NP cue Questionnaire : ___NP PP (blank V)

(11) 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 answer 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 scored significantly higher →

- Storage is reflected in retrieval
- Head-based storage
- Against Independent Storage of phrasal idioms as “big” units

Retrieval experiments in preparation (to assess potentially interfering factors):
   Computerized experiment (Ordentlich in preparation)
   Phrasal idioms in a head-final language
   Literal phrases

(12) 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-selection (on l-selection, see Baltin 1987, O’Grady 1998, Rothstein 1992).

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

(4)  
- a. can't see the forest for the trees  
- b. butter wouldn't melt in X's mouth  
- c. not have a leg to stand on  
- d. the squeaky wheel gets the grease  
- e. what's eating X?

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.

(13) **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).

**Storage hypotheses for clausal idioms**

(14) **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.

(15) **Storage as subentries of the lexical head** of the “extended projection” constituting the clausal idiom (using Grimshaw’s 1991, 2000 term), namely, the verb on a par with VP idioms.

(16) **Independent Storage** as a “big unit” (i.e., not a subentry to another lexical entry)

Independent Storage is a storage strategy motivated independently by idiomatic expressions that have no recognizable internal structure at all, i.e., have neither a functional nor a lexical head, such as in (17), and thus no head-based storage option is available for them.

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

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

(15) and (16) have opposite predictions regarding the distributional properties of clausal
idioms: Only (15) predicts the same pattern of sensitivity to diathesis as attested in verb phrase
idioms (sections 3.2 and 3.3).

3.2 The distribution of unique clausal idioms: the case of verbal passives

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

(16) 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:

(19) a. Rome wasn’t built in one day.
   (They didn’t built Rome in one day: only literal meaning)

   b. may as well be hung for a sheep as (for) a lamb
   (may as well hang X for a sheep as (for) a lamb: only literal meaning)

   c. wouldn’t be caught dead…
   (wouldn’t catch X dead…: only literal meaning)

   d. kurcu me-oto xomer            (fixed tense) (Hebrew)
      made.VPASSIVE of-same material
      ‘share the same characteristics’
      (kerec X me-oto xomer: if at all, only literal, with reference to cutting out cookies
      from dough)

   e. nigzezu maxlafot-av            (fixed tense)
      shear.VPASSIVE hair-one’s
      ‘lost his power/influence’
      (gazezu et maxlafot-av: shear OM hair-his: only literal)

3.3 Evidence from diathesis rigidity: uniqueness across-the-board

Clausal idioms, unlike phrasal ones, fail to exhibit sharing across diatheses: According to our
preliminary investigation they are uniformly unique to their diathesis, irrespective of what this
diathesis is, whether a lexical output, e.g., The squeaky wheel gets the grease vs. One gets the
squeaky wheel the grease, the latter – only literal meaning, or a syntactic output, e.g., (20):
(20) **Clausal idioms: transitive vs. verbal passive: (fixed modal/negation/aux)**

a. can’t make heads or tails of X
b. *Head or tails can’t be made of the proposal. (based on NSW 1994)
c. can’t see the forest for the trees
d. *The forest can’t be seen for the trees.
e. could’ve knocked me over with a feather
f. I could’ve been knocked over with a feather. (only literal meaning)

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

(21) 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.PASS trees, sprinkled. UNACC chips
    (only literal meaning)
c. Kše-xotvim ecim, metizim švavim.
    when-chop.TRANs.IMPERS trees, sprinkled.TRANs chips
    (only literal meaning)

- Sharing of phrasal idioms across diatheses is the result of the links between derivationally related diatheses (entries) in the lexicon.
- Since clausal idioms are stored as units on their own and not as subentries, sharing of idiomatic meaning is not expected.

In sum: Preliminary testing of both of the predictions discussed in 3.2 and 3.3 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 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.

**Rationale:** Subentry storage requires

- obliterating of syntactic structure – less accessible than l-selection
- formation of a new lexical entry – less accessible than addition of subentries?

**Prediction:** If this is on the right track, it predicts that the set of clausal idioms in natural language should be smaller than that of phrasal idioms.

**Preliminary results**

A survey of a sample of 4 letters (2nd to 5th) from Rosenthal Hebrew dictionary reveals:

175 VP idioms vs. 38 clausal ones. (Birger 2014)
● 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:

(22) 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’s (2013)

These intriguing findings (to be discussed in the next presentation) provide support for the validity of the distinction.

Does this distinction correspond to different lexical storage techniques for the two types of phrasal idioms?

4.2 Initial evidence for identical storage

Nondecomposable phrasal idioms behave with regard to being unique vs. shared across diatheses on a par with decomposable phrasal idioms, and in contrast with clausal idioms.

● Nondecomposable phrasal idioms are not rigid regarding diathesis, e.g.,

(23) a. hevi la-olam
    brought to.the-world
    ‘gave birth’

b. ba la-olam
    came to.the-world
    ‘was born’

(24) a. he’ela le-X et laxac ha-dam
    increased.TRANS to-X OM pressure the-blood
    ‘angered X’

b. ala le-X laxac ha-dam
    increased.UNACC to-X pressure the-blood
    ‘angered X’
(25) a. hoci et X me-hakelim
put.out OM X from-the-instruments
‘angered X’

b. yaca me-ha-kelim
went.out.UNACC from-the-instruments
‘became angry’

(26) a. open the door to X
‘enable’

b. the door opened to X
‘become possible’

(27) a. freeze X’s blood
‘horrify/frighten’

b. X’s blood froze
‘get horrified/frightened’

• Nondecomposable phrasal idioms allow passivization; passive nondecomposable phrasal idioms are always ‘shared’, unlike clausal idioms:

(28) 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 (20a))

(29) a. make no bones about X

b. No bones were made about the gravity of the situation.

(30) a. make mincemeat (out) of X

b. Mincemeat will be made (out) of the opposing side.

(31) a. hixnis et X la-tmuna
let.in . OM X to.the-picture
‘involved someone in’

b. huxnas la-tmuna
let.in.VPASS to.the-picture

(32) a. hesir et ha-lot
removed OM the-covering
‘unveiled’

b. husar ha-lot
removed.VPASS the-covering

(33) a. hikši’ax lib-o
hardened heart-his
‘got stubborn’

b. lib-o hukšax
heart-X hardened
The distinction between decomposable and nondecomposable idioms does not seem to correspond to distinct storage strategies (subentry storage vs. independent storage, respectively).

4.3 In progress

- Given that certain nondecomposable idioms but not others can passivize, what governs the passivizability of phrasal idioms?
- To what extent is decomposability relevant as to whether phrasal idioms can be syntactically manipulated and internally modified (Webelhuth, Bargmann & Götze 2014, Sailer 2013)?
- Is the decomposability split observed for phrasal idioms also attested for clausal idioms?

(34) a. can't see the forest for the trees
    b. birds of a feather flock together
    c. the leopard can't change his spots
    d. that dog won't hunt

(35) a. could've knocked me over with a feather
    b. butter wouldn't melt in X's mouth
    c. can take it to the bank
    d. can’t hold a candle to X

- Are internal modification and syntactic manipulation possible for (certain) clausal idioms? If possible for some, is decomposability a relevant factor? For an examination of syntactic manipulation in certain clausal idioms in German, see Richter and Sailer (2009).

Initial findings: Speakers’ judgments range between “impossible modification”, and “intentional playful misquoting of the idiom” or “creative distortion of the idiom”:

(36)a. can't see the (#most obvious/all-important) forest for the (#over-grown/many striking) trees  
    b. Birds of a feather (#that sets them apart from all others) flock together  
    c. The leopard can't change his (#unattractive/most striking) spots.  
    d. That (#tired/uninspired) dog won't hunt.

- Comparison of clausal idioms with hybrid idioms, that is, phrasal idioms containing embedded idiomatic clauses (e.g., bite off more [than X can chew]; know [which side of X’s bread is buttered]; shut the barn door [after the horse has bolted])