Idioms: The Type-Sensitive Storage Model Julia Horvath and Tal Siloni

Introduction Idioms show an inherent 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 by semantic composition. Hence, idioms must be stored in mental representations. Since they involve syntactic structure, the question as to how they are stored is of particular interest and may shed light on the storage component, the lexicon.

Idioms have commonly been sorted according to various dimensions such as decomposability, transparency, size, etc. Based on new empirical data, we suggest a new classification dimension: the lexical vs. functional headedness of the idiom. We put forth and provide support for the hypothesis that the manner of storage of idioms depends on this structural distinction.

Background We define a core set of idiomatic expressions, to serve as a coherent empirical array to be tested:

(1) Idioms (core set): Fixed multilexemic expressions whose meaning is (i) conventionalized-unpredictable and (ii) figurative (metaphoric).

As defined in (1), we focus on multilexemic expressions as they involve special meanings of lexical items relative to one another, and therefore raise questions that the storage of special meanings of single lexical items does not. Further, we take unpredictability and metaphoricity to be basic determinants distinguishing the core set of idioms from other fixed expressions.

Proposal We observe a dichotomy along the headedness dimension and accordingly define the split between phrasal vs. clausal idioms as in (2), illustrated in (3)) and (4), respectively.

- (2) a. Phrasal Idioms are headed by a lexical head. The idioms in (3) are VP idioms.
 - b. Clausal Idioms, not necessarily full clauses, are headed by a clausal functional head diagnosed by a fixed tense, mood, modal, CP-material or presence/absence of negation.
- (3) a. *land on one's feet* 'make a quick recovery'

b. cool one's heels 'wait'

- (4) a. can't see the forest for the trees
 - 'can't see the whole situation clearly because you're focusing at the details'
 - b. not have a leg to stand on
 - 'have no support (for your position)'
 - c. what's eating X?
 - 'What's annoying you?'

We put forward a new model of idiom storage (TSS):

- (5) The Type-Sensitive Storage Model (TSS)
 - 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" units).

Functional elements unlike lexical ones are closed class items, have no descriptive content (Abney 1987), and bear no thematic relation to their complement. Functional elements have often been argued to be stored in a separate lexicon, e.g., Emonds' (2000) 'Syntacticon', or as 'f-morphemes', which are not part of the Encyclopedia (Distributed Morphology). One storage option for clausal idioms would be storage as subentries of their functional head. For instance, *What's eating X?* would be stored as a subentry of its functional head, the interrogative complementizer. This would be storage of entities that have descriptive content in the 'functional lexicon', where entries do not have descriptive content. We discard this

option as incoherent. Empirical evidence based on idiom surveys in English and Hebrew (numerical results will be presented) provides support for independent storage for clausal

Empirical evidence (I) Idioms unique to their diathesis: We have conducted quantitative idiom surveys in English that have reproduced Horvath and Siloni's (2009) results regarding Hebrew, showing that *phrasal* idioms cannot be specific (unique) to the verbal passive, though they can be specific to the transitive, unaccusative, and adjectival passive (e.g., (6)). In contrast, our searches reveal the existence of *clausal* idioms specific to the verbal passive.

- (unique phrasal unaccusative) (6) burst at the seams 'be extremely full or crowded'
 - #burst something at the seams b.
- (7) a. Rome wasn't built in one day. (unique clausal verbal passive) 'Important work takes time.'
 - #(They) didn't build Rome in one day.

This is expected under the TSS model. Phrasal idioms are stored as subentries of the lexical entry of their head (5a). Subentry storage is contingent upon the existence of the (mother) entry in the lexicon. The verbal passive is formed beyond the domain of storage, beyond the lexicon (Baker, Johnson and Roberts 1989, Collins 2004, a.o.). It follows that the verbal passive is not stored, it is not a lexical entry. Hence, it cannot have subentries. Thus, phrasal idioms cannot be unique to the verbal passive because such idioms cannot be stored. The transitive, unaccusative and adjectival passive, in contrast, are formed within the domain of special meanings, in other words, they are stored, and thus can have subentries. Independent storage for clausal idioms (5b) predicts occurrence of unique clausal idioms in the verbal passive, in concert with our findings. Under independent storage, clausal idioms get lexicalized in one piece, following consistent use of the expression in the relevant context. Clausal idioms thus do not require that their subconstituents be represented as entries in the lexicon. They get stored as a whole and can therefore include any diathesis (or other syntactic output). Our model thus predicts the existence of clausal idioms unique to the verbal passive. (II) Diathesis rigidity vs. sharing: Surveys in both English and Hebrew show that phrasal

- idioms can be shared between root-alternates.
- (8) a. burst someone's bubble (phrasal transitive-unaccusative) 'destroy someone's illusion'
 - someone's bubble burst

In contrast, we have not come across clausal idioms shared by root-counterparts in either of the two languages. Under the TSS model, this is expected. Sharing of phrasal idioms between the transitive, unaccusative, or adjectival passive diatheses is the result of the links between root-related entries in the lexicon. Clausal idioms constitute independent entries on their own rather than being subentries of their verbal head, thus sharing of idiomatic meaning across root-related entries is a priori not expected.

(III) Quantity: The set of clausal idioms in natural language is smaller than that of phrasal idioms. 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). We argue that independent storage is more "costly" than the subentry one.

Time permitting, we will discuss the consequences the two distinct storage strategies suggested by the TSS model have regarding the syntactic flexibility of phrasal vs. clausal idioms.